

Universität Bayreuth
Professur Didaktik der Geographie

**Potenziale ausgewählter Reflektiermethoden
für die gezielte Förderung von
Reflektierfähigkeiten im Rahmen eines
Blogprojekts innerhalb einer Bildung für
nachhaltige Entwicklung**

DISSERTATION

zur Erlangung des akademischen Grades
einer Doktorin der Naturwissenschaften (Dr. rer. nat.)
an der Fakultät für Biologie, Chemie und Geowissenschaften
der Universität Bayreuth

vorgelegt von

Anja Hager
aus Bayreuth
Bayreuth, 2020

Die vorliegende Arbeit wurde in der Zeit von Februar 2017 bis Dezember 2019 in Bayreuth am Lehrstuhl Didaktik der Geographie unter Betreuung von Frau Professorin Dr. Gabriele Obermaier angefertigt.

Vollständiger Abdruck der von der Fakultät für Biologie, Chemie und Geowissenschaften der Universität Bayreuth genehmigten Dissertation zur Erlangung des akademischen Grades einer Doktorin der Naturwissenschaften (Dr. rer. nat.).

Dissertation eingereicht am: 12.12.2019

Zulassung durch die Promotionskommission: 08.01.2020

Wissenschaftliches Kolloquium: 05.02.2020

Amtierender Dekan: Prof. Dr. Matthias Breuning

Prüfungsausschuss:

Prof. Dr. Gabriele Obermaier (Gutachterin)

Prof. Dr. Ludwig Haag (Gutachter)

Prof. Dr. Manfred Miosga (Vorsitz)

PD. Dr. Frank Meyer

Das dieser Dissertation zugrundeliegende Projekt wurde an der Universität Bayreuth im Rahmen der gemeinsamen „Qualitätsoffensive Lehrerbildung“ von Bund und Ländern aus Mitteln des Bundesministeriums für Bildung und Forschung unter dem Förderkennzeichen 01JA160 gefördert. Die Verantwortung für den Inhalt dieser Veröffentlichung liegt bei der Autorin.

GEFÖRDERT VOM



Bundesministerium
für Bildung
und Forschung

Danksagung

An dieser Stelle möchte ich mich bei all denjenigen bedanken, ohne die meine Arbeit nicht in dieser Form hätte realisiert werden können.

Mein besonderer Dank gilt meiner Doktormutter Frau Prof. Dr. Gabriele Obermaier für das in mich gesetzte Vertrauen, die Unterstützung bei der Themenwahl und der anschließenden Planung und Umsetzung meines Projekts. Die konstruktiven und stets angenehmen Gespräche gaben mir immer wieder Motivation während der Entstehung meiner Arbeit.

Außerdem möchte ich mich bei meinen Kolleginnen und Kollegen am Lehrstuhl Geographiedidaktik der Universität Bayreuth bedanken: Bei Jonas Braunersreuther für seine engagierte Mitarbeit und Geduld während unserer stundelangen Diskussionen im Rahmen der Auswertung. Bei Kati Barthmann und Christoph Koch für ihre Hilfe und ihren Rat über den gesamten Zeitraum der Entstehung hinweg. Danke auch an alle weiteren Unterstützerinnen und Unterstützer für die kritischen Betrachtungen und differenzierten Anmerkungen und spannenden Diskussionen!

Ein besonderer Dank gilt allen beteiligten Schulen, Schülern & Schülerinnen und Lehrkräften. Ohne ihre oftmals mit hohem Aufwand für mein Projekt eingeräumte Zeit wäre eine hochwertige Datenerhebung nicht möglich gewesen.

Bei meiner Familie möchte ich mich besonders für die Rücksicht und die Unterstützung in den oft aufreibenden Phasen der Arbeit bedanken. Danke auch an meinen Freund Matthias Luft für seine liebevolle und vielseitige Unterstützung. Ich widme diese Arbeit Albert Werthmann, der ihren Abschluss leider nicht mehr miterleben konnte.

Inhaltsverzeichnis

1	Einleitung	1
2	Theoretische Grundlagen.....	3
2.1	Reflektieren und Reflektierfähigkeiten	3
2.1.1	Definitionen und Konzepte	3
2.1.1.1	Konzept des reflexiven Denkens nach Dewey.....	5
2.1.1.2	Konzept des reflexiven Handelns nach Schön	8
2.1.2	Unterschied Reflektieren und kritisches Denken	10
2.2	Bedeutung von Reflektierfähigkeiten im Geographieunterricht	11
2.3	Bedeutung von Reflektierfähigkeiten im Bereich BNE und dem Globalen Lernen	16
2.3.1	Bildung für nachhaltige Entwicklung und Globales Lernen	16
2.3.2	Reflektierfähigkeiten innerhalb der BNE.....	17
2.3.3	Reflektierfähigkeit innerhalb des Orientierungsrahmens für den Lernbereich Globale Entwicklung.....	19
2.4	Messung von Reflektierfähigkeiten.....	22
2.4.1	Schriftliche Erfassung von Reflektierfähigkeiten	22
2.4.2	Modelle zur Bestimmung von Reflektierfähigkeiten	22
2.4.2.1	Brendels Stufenmodell (2017) zur Bestimmung von Reflexionsfähigkeit	25
2.4.2.2	Stufenmodell nach Zimmermann und Welzel (2008) zur Bestimmung von Reflexivitätsfähigkeit	29
2.5	Förderung von Reflektierfähigkeiten.....	33
2.5.1	Theorien zur Förderung von Reflexionsprozessen	33
2.5.2	Methoden zur Förderung von Reflexionsprozessen	36
2.5.2.1	Gibbs Reflexivitätszyklus.....	36
2.5.2.2	Hüttemethode nach de Bono.....	40
2.6	Das digitale Medium Blog	44
2.6.1	Definition und Klassifikation von Medien	44
2.6.2	Blogs als Unterkategorie digitaler Medien	46

2.6.3	Eigenschaften von Blogs.....	47
2.7	Forschungsstand	49
2.7.1	Brendels Dissertationsstudie (2017).....	50
2.7.2	Weitere Studien zu Reflektierfähigkeiten im Rahmen des Geographieunterrichts und des Globalen Lernens	51
2.7.2.1	Forschungsarbeiten zur Förderung von Reflektierfähigkeiten.....	53
2.7.2.2	Forschungsarbeiten mit Reflektierniveaumodellen zur Bestimmung von Reflektierfähigkeiten.....	57
3	Empirische Untersuchung	60
3.1	Forschungsfragen	60
3.2	Übergeordnetes Forschungskonzept und Forschungsdesign.....	61
3.3	Erhebungsverfahren.....	66
3.3.1	Begründung für die Auswahl von Blogs als Haupterhebungsinstrument.....	66
3.3.2	Fragebogen als weiteres Erhebungsinstrument	68
3.3.3	Sampling.....	69
3.3.4	Datenerhebung während des Projekts	72
3.4	Aufbereitung und Auswertung der Daten.....	77
3.4.1	Qualitative Inhaltsanalyse als erster Auswertungsschritt.....	78
3.4.2	Gütekriterien der Inhaltsanalyse.....	96
3.4.3	Quantitative Auswertung als zweiter Auswertungsschritt.....	97
3.4.3.1	Begründung des ausgewählten Auswertungsverfahrens.....	97
3.4.3.2	Hypothesen und Vorgehen bei Auswertung	102
3.5	Ergebnisse.....	104
3.5.1	Einzelauswertungen der Teilnehmenden im Rahmen der qualitativen Inhaltsanalyse	105
3.5.2	Zusammenfassung der Ergebnisse der qualitativen Inhaltsanalyse	112
3.5.3	Ergebnisse der quantitativen Auswertung	139
3.5.3.1	Hypothese 1	139
3.5.3.2	Hypothese 2 und 3.....	143

3.5.3.3	Hypothese 4.....	148
3.5.3.4	Hypothesengenerierender Teil	149
4	Beantwortung der Forschungsfragen	155
5	Diskussion	164
5.1	Erkenntnisse aus der Studie	164
5.2	Grenzen der Studie	167
6	Fazit.....	169
	Literaturverzeichnis.....	171
	Abbildungsverzeichnis	188
	Tabellenverzeichnis	189
	Abkürzungsverzeichnis.....	191
	Anhang	192
	(Eidesstattliche) Versicherungen und Erklärungen	

Summary

Reflection is an ability that plays an important role not only in the school context, but also in everyday situations. It is necessary to enable lifelong learning and it is regarded as a basic requirement for the acquisition of abilities (see Hilzensauer 2008; Rychen 2003).

Nevertheless, only little attention has yet been paid to the selected promotion of the ability to reflect in classroom (see Killus 2009). Based on the findings of Brendel's dissertation (2017) on the determination of students' reflection abilities using weblogs in the context of global learning, this study examines how selected methods are suitable for explicitly promoting reflective skills of adolescents.

The study is realized as part of a blog project with high school seminars at Bavarian grammar schools. For one school year, students are writing texts on predetermined topics in the context of an education for sustainability. They are asked to reflect on the topics themselves and on their own behavioral habits in their articles on the weblog. During the project, the pupils learn two methods at two different times, which they can use to reflect in their articles. The selected reflective methods for this study are Gibbs reflective cycle (1988) and the six thinking hats method according to de Bono (1992). In addition to the blog posts, a questionnaire is completed before and after the intervention in order to be able to discuss the development of reflective skills depending on certain characteristics.

In order to make statements about the potential of the methods for the promotion of reflective skills, this study is adapted to mixed-methods designs. The combination of qualitative and quantitative methods is used in the research design as well as in data collection and -analysis and ensures a multi-perspective examination of the subject of the study (see Tashakkori and Teddlie 1998). First, the posts are evaluated using the qualitative content analysis according to Mayring (2015) in the form of scaling structuring. The texts are classified in view of the content of reflection and reflexivity components and are then classified with two category systems, which are based on existing level models for reflection (see Brendel 2017; Zimmermann and Welzel 2008). Brendel (2017) distinguishes between two types of reflection, in which the focus of reflection is directed differently: At 'reflection' it is reflected on topics and knowledge, at 'reflexivity' it is reflected on actions and behavior. In this study, both sub-skills of reflection will be examined in more detail, as each of them is relevant for the education for sustainability. The data obtained are examined in more detail according to the criteria of score resulting from the percentage distributions to the individual levels within the level models. Other criteria are maximal achieved and most frequently achieved levels. After the first step of evaluation, a quantitative evaluation is done, in which the results of a control group that

does not receive the methods are additionally examined. Variance analyses examine the developments of the criteria that make up the abilities to reflect with respect to significant differences and calculate effect strengths. As part of the statistical evaluation, hypotheses are tested as well as new hypotheses are generated.

The main results of the study are unambiguous: The reflective skills of the students, who learn the two methods, are significantly higher than before the intervention. Both, the ability of reflection and reflexivity, consisting of the three criteria score, maximum and most frequently achieved levels, show enhancements. Gibb's reflexivity cycle significantly improves reflexivity abilities, while de Bono's six thinking hats method leads to a noticeable improvement in reflection abilities. In addition to the above findings, it is also apparent that the potential of aids to reflect is not the same for all pupils, but that there are gender and individual differences. Altogether, the chosen methods are a useful instrument to explicitly promote reflective abilities in the context of sustainable development education based on the insights made in this study. Teachers can therefore use them as an effective instrument for their teaching.

Zusammenfassung

Reflektieren ist eine Fähigkeit, die nicht nur im schulischen Kontext, sondern auch in alltäglichen Situationen eine wichtige Rolle spielt. Sie ist erforderlich, um lebenslanges Lernen zu ermöglichen und gilt als eine Grundvoraussetzung für den Kompetenzerwerb (vgl. Hilzensauer 2008). Dennoch wird der gezielten Förderung der Fähigkeit des Reflektierens bisher noch wenig Aufmerksamkeit im Unterricht geschenkt (vgl. Killus 2009). Auf Grundlage der Erkenntnisse aus Brendels Dissertationsstudie (2017) über die Bestimmung von Reflexionsfähigkeiten von Schülerinnen und Schülern mithilfe von Weblogs im Kontext Globalen Lernens wird in dieser Studie der Frage nachgegangen, inwieweit ausgewählte Reflektiermethoden geeignet sind, um explizit Reflektierfähigkeiten von Jugendlichen zu fördern.

Die Studie findet im Rahmen eines Blogprojekts mit Oberstufenseminaren an bayerischen Gymnasien statt. Über ein Schuljahr hinweg verfassen die Schülerinnen und Schüler Texte zu vorgegebenen Themen, die sie im Rahmen einer Bildung für nachhaltige Entwicklung im Unterricht behandeln. Sie werden aufgefordert, über die Themen selbst und über eigene Verhaltensgewohnheiten diesbezüglich auf einem Weblog zu reflektieren. Während des Projekts erlernen die Testpersonen zu zwei unterschiedlichen Zeitpunkten zwei Reflektiermethoden, an denen sie sich für das Verfassen ihrer Beiträge orientieren sollen. Die gewählten Reflektiermethoden für diese Studie sind Gibbs *reflective cycle* (1988) und die *six thinking hats* Methode nach de Bono (1992). Zusätzlich zu den Blogbeiträgen wird jeweils vor und nach dem Projekt ein Fragebogen ausgefüllt, um näher auf die Entwicklung der Reflektierfähigkeiten in Abhängigkeit bestimmter Merkmale eingehen zu können.

Um Aussagen über die Potenziale der Reflektiermethoden für die Förderung von Reflektierfähigkeiten treffen zu können, wird sich an *mixed-methods*-Designs orientiert. Die Kombination qualitativer und quantitativer Methoden findet sowohl im Forschungsdesign, als auch in der Datenerhebung und -auswertung statt und gewährleistet eine mehrperspektivische Analyse des Untersuchungsgegenstands (vgl. Tashakkori und Teddlie 1998). Zunächst werden die Beiträge mithilfe der qualitativen Inhaltsanalyse nach Mayring (2015) in Form der skalierenden Strukturierung ausgewertet. Die Texte werden hinsichtlich enthaltener Reflexions- und Reflexivitätsanteile eingeordnet und anschließend mit zwei Kategoriensystemen, die sich an bestehende Niveaumodellen fürs Reflektieren orientieren (vgl. Brendel 2017; Zimmermann und Welzel 2008), kodiert. Brendel (2017) unterscheidet zwischen zwei Reflektierarten, bei denen der Fokus des Reflektierens unterschiedlich gerichtet ist: Bei der Reflexion wird über Themen und Wissen reflektiert, bei der Reflexivität über Handlungen. In dieser Studie sollen beide Teilfähigkeiten der Reflektierfähigkeit insgesamt genauer untersucht werden, da

jede von ihnen relevant innerhalb einer Bildung für nachhaltige Entwicklung ist. Die gewonnenen Daten werden nach den Kriterien Score, der sich aus den prozentualen Verteilungen auf die einzelnen Stufen innerhalb der Niveaumodelle ergibt, maximal erreichte und am häufigsten erreichte Stufen genauer untersucht. Nach dem ersten Auswertungsschritt folgt eine quantitative Auswertung, bei der zusätzlich die Ergebnisse einer Kontrollgruppe herangezogen werden, die keine Reflektiermethoden erhält. Durch Varianzanalysen werden die Entwicklungen der Kriterien, aus denen sich die Reflektierfähigkeiten zusammensetzen, bezüglich signifikanter Unterschiede untersucht und Effektstärken berechnet. Im Rahmen der statistischen Auswertung werden sowohl Hypothesen getestet als auch neue Hypothesen generiert.

Die Hauptergebnisse der Studie sind eindeutig: Die Reflektierfähigkeiten der Schülerinnen und Schüler, die die beiden Reflektiermethoden erlernen, sind nach dem Blogprojekt deutlich höher als zuvor. Sowohl die Reflexions- als auch die Reflexivitätsfähigkeiten, bestehend aus den jeweils drei Kriterien Score, maximal und am häufigsten erreichte Stufen, zeigen Verbesserungen. Gibbs Reflexivitätszyklus bewirkt eine deutliche Verbesserung der Reflexivitätsfähigkeiten, während die Hütemethode von de Bono zu einer erkennbaren Verbesserung der Reflexionsfähigkeiten führt. Neben den genannten Feststellungen zeigt sich ebenfalls, dass die Potenziale der Reflektiermethoden nicht bei allen Probandinnen und Probanden gleich ausfallen, sondern dass geschlechtsspezifische und individuelle Unterschiede bestehen. Insgesamt kommt die Studie zu dem Ergebnis, dass sich die gewählten Methoden für den Einsatz im Unterricht zur expliziten Förderung von Reflektierfähigkeiten im Rahmen einer Bildung für nachhaltige Entwicklung eignen und von Lehrkräften als Instrument verwendet werden können.

1 Einleitung

Reflektieren – ein Begriff, der in vielen Lebensbereichen gerne verwendet wird, auch wenn dessen genaue Bedeutung häufig ungeklärt bleibt. Sucht man online nach dem Begriff, so stellt man fest, dass er in unterschiedlichsten Kontexten millionenfach verwendet wird. Die Mächtigkeit des Begriffs ist folglich unumstritten. Auch im schulischen Kontext spielt Reflektieren hinsichtlich der zu erwerbenden Kompetenzen eine wichtige Rolle (vgl. Hilzensauer 2008; Rychen 2003). Von Schülerinnen und Schülern wird erwartet, dass sie reflexive Fähigkeiten besitzen, doch kaum bis selten wird diese Fähigkeit als eigener Unterrichtsgegenstand behandelt oder explizit eingeübt (vgl. Killus 2009). Im Rahmen einer Bildung für nachhaltige Entwicklung und im Geographieunterricht kann die Fähigkeit des Reflektierens als Grundlage für einige Kompetenzen, wie beispielsweise den Kompetenzbereich Bewerten/Beurteilen oder im Rahmen der Gestaltungskompetenz gesehen werden (vgl. Kapitel 2.2 und Kapitel 2.3). Der Begriff des Reflektierens wird hier explizit zur Beschreibung der Kompetenzen verwendet, jedoch wird er selbst hier nicht genauer definiert oder Hinweise für die Verwendung im Unterricht gegeben.

Ausgehend von den beiden Eigenschaften von Reflektieren im schulischen Bereich, nämlich der Wichtigkeit für den Kompetenzerwerb und der nur seltenen, expliziten Förderung von Reflektierfähigkeiten, liegt der Fokus dieser Studie auf der Untersuchung der Potenziale ausgewählter Methoden auf die schriftlich gezeigten Reflektierfähigkeiten von Schülerinnen und Schülern. Während eines Blogprojekts, bei dem Jugendliche über Themen, die für eine Bildung für nachhaltige Entwicklung relevant sind und über ihr Verhalten diesbezüglich schreiben, erlernen sie zwei Reflektiermethoden, an denen sie sich fürs Reflektieren orientieren können. Dadurch sollen Erkenntnisse über die explizite Förderung von Reflektierfähigkeiten gewonnen werden, von denen aktuell nur wenige vorliegen (vgl. Brendel 2017, S. 277). Im Idealfall sollen die für diese Studie entwickelten und getesteten Methoden als Einsatzmöglichkeiten für Lehrkräfte dienen, um das Reflektieren bewusst mit ihren Schülerinnen und Schülern üben zu können. Eine Untersuchung, die Potenziale von Reflektiermethoden auf die entsprechende Fähigkeit zum Forschungsschwerpunkt wählt, setzt eine gründliche Auseinandersetzung mit Theorien zum Reflektieren voraus, ebenfalls müssen die Bereiche der Messung und der Förderung von Reflektierfähigkeiten ausreichend berücksichtigt werden. Als theoretischer Rahmen für die Studie werden deshalb folgende Aspekte aufgegriffen: Zunächst wird ausführlich auf den Begriff des Reflektierens und die Unterscheidung von Reflexion und Reflexivität eingegangen (vgl. Kapitel 2.1.1). Anschließend wird geklärt, welche Bedeutung das Reflektieren im Geographieunterricht und im Lernbereich Globales Lernen trägt

(vgl. Kapitel 2.2 und Kapitel 2.3). Kapitel 2.4 setzt sich mit Modellen zur Messung von Reflektierfähigkeiten auseinander und stellt die beiden Modelle, die als Grundlage für die Bestimmung der Reflektierfähigkeiten in dieser Studie eingesetzt werden, genauer vor. Daran anschließend werden Theorien zur Förderung von Reflektierfähigkeiten und die Methoden, die im Rahmen dieser Studie als Reflektierhilfen¹ dienen sollen, behandelt (vgl. Kapitel 2.5). In einem weiteren Punkt wird genauer auf das Medium Blog eingegangen, da Blogs für die Erhebung der Daten in dieser Studie eingesetzt werden (vgl. Kapitel 2.6). Ein Überblick über den Forschungsstand im Bereich von Reflektierfähigkeiten im Rahmen des Geographieunterrichts und einer Bildung für nachhaltige Entwicklung, sowie die Messung und Förderung dieser schließt den theoretischen Rahmen ab (vgl. Kapitel 2.7). Der empirische Teil der Arbeit beginnt mit einer Vorstellung der Teilforschungsfragen, um die Hauptforschungsfrage beantworten zu können (vgl. Kapitel 3.1), die lautet: Welche Potenziale bieten ausgewählte Reflektiermethoden für die gezielte Förderung von Reflektierfähigkeiten bestehend aus Reflexions- und Reflexivitätsfähigkeiten von Jugendlichen beim Bloggen im Rahmen einer Bildung für nachhaltige Entwicklung?

Weiterhin unterteilt sich der empirische Teil der Arbeit in die Beschreibung und Begründung des hier verwendeten *mixed-methods*-Forschungsdesigns, die Projektbeschreibung inklusive Datenerhebung und die anschließende Datenauswertung. Letzteres wird in zwei Teilabschnitte unterteilt: Zunächst wird eine qualitative Inhaltsanalyse nach Mayring (2015) durchgeführt, anschließend folgt eine quantitative Auswertung in Form von Varianzanalysen. Im Ergebnisteil werden die Resultate der qualitativen und quantitativen Auswertung zusammengefasst und dadurch Antworten auf die (Teil-)Forschungsfrage(n) gegeben (vgl. Kapitel 4). Die darauffolgende Diskussion dient dazu, Erkenntnisse für die Implementierung von Förderungsmöglichkeiten von Reflektierfähigkeiten im Unterricht zu diskutieren (vgl. Kapitel 5). Außerdem werden vor dem abschließenden Fazit die Grenzen der Studie dargestellt.

¹ Als Synonym für Reflektiermethode wird im Rahmen dieser Studie ebenfalls der Begriff *Reflektierhilfe* verwendet.

2 Theoretische Grundlagen

In diesem Kapitel, dem theoretischen Teil, werden die Grundlagen für die hier vorliegende Forschungsarbeit vorgestellt. Ein Überblick über den aktuellen Forschungsstand innerhalb dieses Gebiets rundet das Kapitel ab.

2.1 Reflektieren und Reflektierfähigkeiten²

Die Fähigkeit des Reflektierens ist in vielen Lebensbereichen von Bedeutung. Sowohl im Alltag als auch in der Literatur wird der Begriff *Reflektieren* häufig verwendet. Was man darunter aber genau versteht, bleibt meist ungewiss (vgl. Reinmann 2005; Smyth 1992). Reflektieren kann sich auf unterschiedliche Aspekte beziehen, weshalb ungenaue Begriffsklärungen schnell zu Missverständnissen führen können. Aufgrund existierender Verständnisschwierigkeiten, die häufig mit den Begriffen *Reflektieren*, *Reflexion* und *Reflexivität* auftreten und aufgrund der Tatsache, dass die hier vorliegende Studie Reflektierfähigkeiten als Untersuchungsinteresse hat, ist es unerlässlich, unterschiedliche Konzepte und -arten des Reflektierens genauer vorzustellen. Im Folgenden wird deshalb die Problematik der ungenauen Begriffsdefinition aufgegriffen und eine Arbeitsdefinition für den Begriff *Reflektieren* vorgestellt, die für diese Studie als Grundlage dient. Dazu wird zwischen *Reflexion* und *Reflexivität* unterschieden, die als zwei Teilbereiche des Reflektierens gesehen werden und die auf zwei unterschiedlichen Konzepten beruhen, die in Kapitel 2.1.1.1 und 2.1.1.2 noch ausführlich behandelt werden.

2.1.1 Definitionen und Konzepte

Reflection can mean all things to all people... it is used as a kind of umbrella or canopy term to signify something that is good or desirable... everybody has his or her own (usually undisclosed) interpretation of what reflection means, and this interpretation is used as the basis for trumpeting the virtues of reflection in a way that makes it sound as virtuous as motherhood. (Smyth 1992, S. 285)

Smyth Zitat verdeutlicht das eben beschriebene Problem, dass es keine generell gültige Definition für den Begriff *Reflektieren* gibt und dass das Begriffsverständnis je nach Disziplin unterschiedlich aufgefasst werden kann. Dennoch wird der Begriff sowohl im wissenschaftlichen als auch privaten Sprachgebrauch häufig verwendet, mit der Annahme, jeder wisse, was genau darunter zu verstehen sei. Auch in der pädagogischen, psychologischen und geographie-didaktischen Literatur unterliegen Begriffe wie *Reflektieren*, *Reflexion* oder *Reflexivität* keiner einheitlichen Begriffsklärung und werden zum Teil sogar ohne konkrete Definition verwendet (vgl. Abou Baker El-Dib 2007). So heißt es in den Bildungsstandards im Fach Geographie für

² In dieser Studie wird bewusst der Begriff *Reflektierfähigkeit* verwendet, um die allgemeine Fähigkeit des Reflektierens zu beschreiben. Unter dem gängigeren Begriff *Reflexionsfähigkeit* wird im Rahmen dieser Studie nur ein Teilbereich davon verstanden (vgl. 2.1.1.1).

den Mittleren Schulabschluss, dass „eine[] reflektierte[], ethisch begründete[] und verantwortungsbewusste[] raumbezogene[] Handlungsfähigkeit“ (DGfG 2014, S. 8)³ eines der Leitziele des Faches sei. Auf das Verständnis dieser Fähigkeit wird dabei jedoch nicht genauer eingegangen und somit bleibt die Frage offen, was unter den Begriff *reflektiert* in diesem Kontext eigentlich verstanden wird. Dadurch besteht die Gefahr, dass jeder ausgehend seines Vorwissens Reflektieren unterschiedlich versteht und ihm eigene Charaktermerkmale zuschreibt. Doch nicht nur im deutschsprachigen Raum, sondern auch im internationalen Diskurs und vor allem im Forschungsbereich der Lehramtsausbildung werden Begriffe wie *reflection*, *critical reflection*, *reflective action*, *reflective practitioner* oder *reflective thinking* häufig verwendet ohne angemessen definiert zu werden (vgl. Hatton und Smith 1995, S. 33). Da es im Englischen viele Termini für den Vorgang des Reflektierens gibt und dieser im Deutschen hauptsächlich mit dem Begriff der *Reflexion* umschrieben wird, bedarf es einer genauen Betrachtung einzelner Definitionsversuche und dahinterliegender Konzepte, um adäquat mit den Begriffen umgehen zu können. Auch im Rahmen der Recherchearbeit stößt man immer wieder auf das Problem der vielfältigen und teilweise ungenauen Begriffsverwendungen für den Vorgang des Reflektierens im Sinne von Nachdenken. Um der Gefahr von Verständnisproblemen entgegenzuwirken ist deshalb wichtig, vor Verwendung des Begriffs klarzustellen, was man als Autorin bzw. Autor selbst darunter versteht und auf welchen Grundlagen dieses Verständnis aufbaut. Im Folgenden soll deshalb eine Arbeitsdefinition zum Begriff *Reflektieren* gegeben werden, bevor anschließend die dahinterliegenden Konzepte genauer vorgestellt werden.

Fast ausnahmslos werden für die Definitionsbestimmung von *Reflektieren* die Konzepte von John Dewey (1933) und Donald Schön (1983) als Grundlage genommen (vgl. Abels 2011, S. 49). Auch Brendel (2017) orientiert sich an den beiden Konzepten, um das Reflektieren in zwei Arten, nämlich der Reflexion und der Reflexivität unterscheiden zu können. Für diese Studie werden ebenfalls die beiden Bezeichnungen *Reflexion* und *Reflexivität* nach dem Verständnis von Brendel (2017) verwendet. Sie definiert die beiden Unterformen des Reflektierens wie folgt: „Reflexion bezeichnet ein Nachdenken über Unterrichtsinhalte, Überzeugungen oder Vorstellungen [...]. Reflexivität bezeichnet ein Nachdenken über eigene Handlungen oder Handlungsalternativen“ (ebd., S. 19). Als Oberbegriff und somit zusammenfassender Prozess von Reflexion und Reflexivität wird in dieser Studie *Reflektieren* verwendet, der in Anlehnung an Brendel (2017) und Schmal (2017, S. 25) folgendermaßen definiert wird:

Reflektieren bedeutet Sachverhalte, Wissen, eigene und fremde Verhaltensweisen und Einstellungen kritisch zu hinterfragen und diese daraufhin zu überprüfen, warum sie ihre Gültigkeit und ihren Wert besitzen (oder auch nicht).

³ DGfG steht für Deutsche Gesellschaft für Geographie

Unter Reflektieren kann also ein tiefgreifendes Nachdenken über bestimmte Themen und Handlungen verstanden werden.

In den beiden nachfolgenden Unterkapiteln werden die beiden Konzepte von Dewey (1933) und Schön (1983) ausführlicher vorgestellt, da diese gemeinsam den Rahmen für die hier verwendete Definition von Reflektieren formen.

2.1.1.1 Konzept des reflexiven Denkens nach Dewey

Reflexion nach dem Verständnis von Brendel (2017) beruht auf Deweys Konzept des reflexiven Denkens, das er bereits 1910 in seinem Werk *How we think* vorstellte. Obwohl Dewey durchgehend in Veröffentlichungen über das Thema Reflektieren erwähnt wird, fehlt in den meisten Werken eine intensive Auseinandersetzung mit Deweys Verständnis von Reflektieren (vgl. Rodgers 2002, S. 843). Deshalb ist es wichtig, das Grundverständnis in dieser Arbeit detailliert vorzustellen, damit anschließende Ableitungen und Interpretationen nachvollziehbar sind.

Rodgers arbeitet vier Kriterien aus Deweys Werken heraus, die seine Urgedanken zugänglich und transparent darstellen (vgl. ebd., S. 845):⁴

- Reflektieren als bedeutungsgebender Prozess
- Reflektieren als gründliche Denkweise
- Reflektieren in Gemeinschaft
- Reflektieren als Reihe von Einstellungen

Nachfolgend wird auf diese vier Kriterien mithilfe von Originalausschnitten aus Deweys Werken genauer eingegangen, damit Deweys Grundverständnis von Reflektieren deutlich wird. Dewey untersucht in seinen Werken und vor allem im Urwerk *How we think* (1910) das Ziel von Erziehung. Er sieht darin die Evolution zu einer demokratischen Gesellschaft. Erziehung definiert er dabei mit folgenden Worten: „[R]econstruction or reorganization of experience which adds to the meaning of experience, and which increases [one's] ability to direct the course of subsequent experience“ (Dewey 2009, S. 50). Nach Rodgers (2002, 846 ff.) Interpretation wird Erziehung in dieser Definition eher als Verb denn als Nomen verstanden und genau dieser Prozess wird durch Reflektieren ermöglicht. Um Reflektieren zu können, benötigt man folglich Erlebnisse bzw. Erfahrungen, die nach Dewey jedoch sehr umfassend ausfallen können. Zu Erlebnissen zählt er neben der physischen Teilnahme an Ereignissen auch kogni-

⁴ Rodgers Artikel ist in englischer Sprache verfasst, die Forscherin hat die definierten Kriterien ins Deutsche übersetzt. Die Originalbezeichnungen lauten: „reflection as a meaning making process“, „reflection as a rigorous way of thinking“, „reflection in community“, „reflection as a set of attitudes“ (Rodgers 2002, S. 845).

tive Erfahrungen, wie beispielsweise die Interaktion mit Objekten, die in der Vorstellung konstruiert werden, das Lesen eines Buchs oder die Diskussion mit anderen. Wichtig dabei ist immer die Interaktion zwischen der Person und ihrem Environment.⁵ Environment ist dabei „whatever conditions interact with personal needs, desires, purposes, and capacities to create the experience which is had“ (Dewey 1938, S. 43). Folglich wird eine Erfahrung erst zu einer wirklichen Erfahrung, wenn eine Interaktion zwischen dem Selbst und dem Environment in Form von beispielsweise einer anderen Person, der materiellen Welt, der natürlichen Welt oder einer Idee stattfindet. Als zweite Voraussetzung für Erfahrung gilt Kontinuität, die für Dewey zwei Ausprägungen aufweist: Kontinuität einer Erfahrung durch das Erneuern der sozialen Gruppe und Kontinuität in Form eines Schema-Aufbaus für nachfolgende Erfahrungen (vgl. ebd., S. 39 ff.). Ersteres wird durch Erziehung als Instrument erreicht und beispielsweise durch Technologie, Gesetze oder zivilisierte Verhaltensweisen ausgedrückt. Letzteres drückt aus, dass Wissen und Fähigkeiten, die in einer Situation erlernt wurden, als Instrument dienen, um künftige Situationen verstehen und mit ihnen umgehen zu können. Dieser Vorgang existiert so lange wie Leben und Lernen stattfinden (vgl. ebd., S. 44). Erfahrungen können nach den Worten von Dewey „mis-educative“ (ebd., S. 25) sein, wenn sie zu unüberlegten Routinehandlungen überleiten oder „educative“ (ebd.), wenn sie in eine konstruktive Richtung führen und daraus eine intelligente Handlung entsteht. Damit Erfahrungen sinnvoll und nachhaltig werden, wird die Fähigkeit des Reflektierens benötigt. Rodgers (2002) fasst die Bedeutung des Reflektierens mit Fokus auf die Erfahrung sehr gut zusammen, weshalb ich seine Worte abschließend für das erste Kriterium aufführen möchte:

The function of reflection is to make meaning: to formulate the „relationships and continuities“ among the elements of an experience, between that experience and other experiences, between that experience and the knowledge that one carries, and between that knowledge and the knowledge produced by thinkers other than oneself. (ebd., S. 848)

Das zweite Kriterium handelt vom Reflektieren als gründliche Denkweise, die sich von anderen Denkformen unterscheidet. Dewey (1933) unterscheidet zwischen vier verschiedenen Arten des Denkens. Die erste Ausprägung des Denkens ist dabei die einfachste Art, zu der alle Gedanken zählen, die einem durch den Kopf gehen. Die vierte Ausprägung ist nach Dewey die höchstmögliche Art des Denkens und umfasst die Aspekte des Hinterfragens des Ursprungs und der Folgen bestimmter Überzeugungen, das Schlussfolgern und das kritische Prüfen bestehender Glaubensgrundsätze (vgl. ebd.). Nur diese höchste Ausprägung des Denkens ist für Dewey ein Vorgang des Reflektierens, der einen bildenden Wert besitzt. Die anderen drei Arten, die er als „stream of consciousness“, „invention“ und „belief“ (Dewey 1933,

⁵ Dewey benutzt ebenfalls den Begriff *environment*. Da Übersetzungen die Bedeutung des Begriffs unter der Verwendung von Dewey nicht exakt einfangen, wird deshalb auch im Deutschen die Bezeichnung *Environment* genommen.

S. 4 ff.) bezeichnet, sind reine Denkprozesse, die alleine betrachtet keinen Mehrwert liefern. Dewey definiert die vierte Art des Denkens mit folgenden Worten: „Active, persistent, and careful consideration of any belief or supposed form of knowledge in the light of the grounds that support it, and the further conclusions to which it tends“ (ebd., S. 9). Um reflektierend nachzudenken ist es demnach nötig, sich selbstkritisch zu betrachten und seinen eigenen Gedanken mit Skepsis zu begegnen. Der Unterschied zwischen reflektiertem und nicht-reflektiertem Denken ist der, dass bei ersterem eine überlegte Handlung resultiert, während bei letzterem eine impulsive Handlung folgt (vgl. Dewey 1933, 9 ff.). Dies entspricht dem eben vorgestellten ersten Kriterium, wodurch eine wechselseitige Beziehung beider Kriterien ersichtlich wird.

Der Prozess des Reflektierens wird durch Perplexität ausgelöst. Diese wird erzeugt, wenn eine Person einer Situation – mental oder physisch – begegnet, deren voller Charakter noch nicht bestimmbar ist und die Bedeutung der Erfahrung noch nicht vollständig erfasst werden kann (vgl. ebd., S. 10). Mit dem Vorhandensein der motivationalen Komponente der Kuriosität und dem Wunsch nach Lösung des Ungleichgewichts wird der Prozess des Reflektierens initiiert. Dieser Prozess besteht aus mehreren Teilen. Die Anzahl der Teilprozesse unterscheidet sich in Deweys Werken, jedoch lassen sich die Abschnitte nach Rodgers (2002) in vier Phasen zusammenfassen:

- Auftreten einer Erfahrung
- Beschreibung der Erfahrung
- Analyse der Erfahrung
- Intelligente Handlung/Erprobung⁶

In diesem Zusammenhang möchte ich auf einen Aspekt genauer eingehen: Das Unterschlagen der vierten Phase. Nach Rodgers (2002, S. 855 f.) vergessen viele Autorinnen und Autoren, die über Reflektieren schreiben, die letzte Phase, als Schritt des Reflektierens. Diese besagt, dass der Prozess des Reflektierens eine Handlung beinhalten muss. Diese Handlung kann zunächst auch nur ein Ausprobieren von Ideen sein, die aus der Analyse heraus entstanden sind. In der Erprobung aufgestellter Theorien können neue Fragen, Probleme oder Ideen entstehen, die nachfolgend wieder den Prozess des Reflektierens initiieren. Somit ist Reflektieren ein zyklischer Prozess: Eine Erfahrung, die durch Perplexität gründliches Denken mit sich bringt, führt zu einer nächsten Erfahrung, die wiederum die gleichen Folgen mitbringt (vgl. ebd.).

Das dritte, wichtige Kriterium zum Verständnis des Begriffs des Reflektierens nach Dewey besagt, dass Reflektieren in Gemeinschaft vollzogen werden muss. Dewey (1933, S. 6) schreibt, dass Erfahrungen formuliert werden müssen, um kommuniziert werden zu können.

⁶ Eine genaue Ausführung der einzelnen Phasen kann in Rodgers Artikel *Defining Reflection: Another Look at John Dewey and Reflective Thinking* (2002) nachgelesen werden.

Dabei spielt die Sprache eine wichtige Rolle. Erfahrungen müssen mit Erfahrungen anderer verglichen werden, um sie analysieren zu können. Um Erfahrungen anderer überhaupt kennenlernen zu können, muss das reflektierende Denken nach außen getragen werden. Erst so können Schwächen und Stärken des Denkprozesses beim eigenen Reflektierprozess berücksichtigt werden (vgl. ebd.).

Als letztes Kriterium wird das Reflektieren als Reihe von Einstellungen aufgeführt. Dewey (1933, S. 278) betont, dass Einstellungen und Gefühle bedeutend beim reflexiven Denken und ausschlaggebend dafür sind, in welche Richtung der Prozess verläuft. Ein guter Denker bringt die Eigenschaft mit sich, sich seinen Einstellungen und Emotionen bewusst zu sein und diese entsprechend dem eigenen Vorteil zunutze zu machen.

Zusammenfassend lässt sich festhalten, dass Reflektieren ein iterativer, nach vorne gerichteter Prozess ist, der sowohl Praxis als auch Theorie beinhaltet, die sich gegenseitig abwechseln und beeinflussen. Reflektieren ist ein Denkprozess, der strukturiert und systematisch abläuft und zu Bedeutung führt und sich damit von anderen Denkprozessen unterscheidet (vgl. Dewey 1933). Erfahrungen, die Perplexität auslösen, führen zum Reflektierprozess. Dabei können Erfahrungen sowohl mental als auch physisch stattfinden. Die Erkenntnisse, die sich aus dem Reflektieren ergeben, werden in neuen Erfahrungen getestet und führen wieder zu neuen Denkprozessen und Erkenntnissen. Nicht zu vernachlässigen sind dabei Emotionen und Einstellungen, die die Richtung des Reflektierens beeinflussen können (vgl. Rodgers 2002, S. 863). Reflektieren nach dem Verständnis von Dewey kann folglich sowohl auf Inhalte als auch auf Handlungen gerichtet sein, je nachdem, ob eine Erfahrung geistig oder real gemacht wird. Allerdings liegt der Fokus seiner Überlegungen eindeutig auf dem reflexiven Denken, wie es bereits der Titel *How we think* (1910) postuliert. Auch Brendel (2017) ist der Meinung, dass „der Fokus auf Wissenselementen und die Notwendigkeit des Hinterfragens“ (ebd., S. 17) liegt. Deweys Überlegungen können deshalb als Grundlage für die Definition des Begriffs *Reflexion* verstanden werden. Die Ideen des reflexiven Denkens schafften die Grundlage für Schöns Konzept der *reflective practice* (1983), das sozusagen als Weiterentwicklung des eben beschriebenen Konzepts gesehen werden kann und in dieser Studie als Grundlage für Reflexivität dient.

2.1.1.2 Konzept des reflexiven Handelns nach Schön

Neben Dewey beschäftigt sich auch Donald Schön (1983) ausführlich mit dem Konzept des Reflektierens. Unter seinem Konzept des *reflective practitioners* versteht er Reflektieren als kritisches Nachdenken über Handlungen (vgl. ebd.). Das eigene Handeln zu reflektieren ist seiner Ansicht nach wichtig, da professionelle Handlung durch die Betonung einer positivistischen „Technical Rationality“ (ebd., S. 21) auf starres Anwenden wissenschaftlicher Theorien und Methoden und auf instrumentelles Problemlösen reduziert wird, während die Gesellschaft

gleichzeitig immer weniger Vertrauen in die Relevanz und Angemessenheit von Expertenwissen hat. Durch die Orientierung auf Vorgegebenes wird Theorien mehr Bedeutung zugesprochen als deren Anwendung. Dadurch kommt es zu einer Trennung von Theorie und Praxis, die Schön durch den Vorgang des Reflektierens beheben möchte (vgl. ebd., S. 37). Für ihn ist eine Person reflektiert, wenn sie sich den zugrundeliegenden Einstellungen für bestimmte Entscheidungen und Handlungen bewusst ist und somit nicht nur Wissen anwendet, sondern auch hinterfragt, warum man dies als richtig ansieht (vgl. Copeland et al. 1993, S. 352). Seine Vorstellung von Reflexion, die er als *reflective practice* (Schön 1983) bezeichnet, wird im Deutschen oftmals mit dem Begriff *Reflexivität* umschrieben. So definiert Wirtz (2014) Reflexivität aus kognitionspsychologischer Sicht als „Tendenz zum überlegten, besonnen Handeln im Ggs. zur Impulsivität“ (ebd., S. 1399). Bock (2001) behandelt das Thema Reflexivität in einem Artikel und schreibt dazu:

Reflexivität meint also nicht Reflexion, sondern die Möglichkeit, zwischen Antrieb und Handlung einen Hiatus einzuschalten (vgl. Zdarzil 1978), sich auf sich selbst zu beziehen und sich als Urheber und Grund von Handlungen zu erfahren. Damit verbunden ist eine Wertung des eigenen Tuns auf verschiedenen Ebenen als misslungen oder gelungen, erfolgreich oder vergeblich, angemessen oder unangemessen, bis hin zu gut oder böse. So eröffnet die Reflexivität auch eine ethische und sittliche Dimension. (ebd., S. 115)

Schön (1983, S. 61) unterscheidet zwischen zwei Arten von Reflexivitätsprozessen: *Reflection-in-action* bezieht sich auf das intuitive Wissen, wie wir Handlungen ausführen. Hierbei existiert eine starke Beziehung zwischen Reflektier- und Aktionskomponenten, da aus dem Reflektieren über die eigene Handlung die zugrundeliegende praktische Theorie herausgearbeitet wird. *Reflection-on-action* bezeichnet hingegen ein Nachdenken über bereits stattgefundenen Aktionen und das Herausfinden der Gründe, warum sie so stattgefunden haben. Daraus können Alternativen für künftige Handlungen abgeleitet werden (vgl. ebd.).

Das Konzept der *reflective practice* nach Schön (1983) findet vor allem im angloamerikanischen Raum Einzug in den Bildungsbereich (vgl. Kinsella 2007). Schön (1992) selbst sieht sein Konzept als „[his] version of Dewey's 'Reflective Thought'“ (ebd., S. 123), das den Fokus mehr auf konkretes Handeln legt. Der Unterschied zwischen Deweys und Schöns Konzepten des Reflektierens liegt in den Ausgangspunkten. Während bei Schön das praktische Handeln im Fokus steht, muss der Auslöser bei Dewey nicht unbedingt eine beobachtbare Situation sein. Es kann sich dabei auch um eine intellektuelle Fragestellung oder ein Gedankenexperiment handeln (vgl. King und Kitchener 1994, S. 6). Für die hier vorliegende Studie dient Schöns Konzept als Grundlage für die Reflexivität, die in dieser Studie als Synonym für das

tiefgreifende Nachdenken über eigenes Verhalten oder Handlungen im Nachhinein⁷ verstanden wird. Reflexion hingegen wird als Synonym für ein tiefgreifendes Nachdenken über Inhalte, Wissen oder Vorstellungen nach dem Verständnis von Dewey (1910; 1933) verstanden.

2.1.2 Unterschied Reflektieren und kritisches Denken

Nachdem die Unterscheidung zwischen Reflexion und Reflexivität vorgenommen wurde, ist es ebenfalls wichtig, den allgemeinen Begriff des Reflektierens vom kritischen Denken zu differenzieren. Für das kritische Denken gibt es wie für das Reflektieren keine einheitlich akzeptierte Definition (vgl. Hilscher 2017). Es lassen sich jedoch Eigenschaften identifizieren, die dem kritischen Denken bei Verwendung der meisten Autorinnen und Autoren als Fertigkeit zugeschrieben werden. Dazu zählen nach Facione (1990, S. 5 ff.) die Interpretation, Analyse, Bewertung und Schlussfolgerung, die zu einer bewussten Urteilsbildung führen. Kritisches Denken bezeichnet folglich den Prozess, der in einer rationalen Urteilsbildung endet. In der Literatur werden Reflektieren und kritisches Denken als ähnliche Fertigkeiten gesehen, da ihre Eigenschaften nah beieinanderliegen. Je nach Disziplin beruhen sie jedoch auf leicht unterschiedlichen Verständnissen (vgl. Moon 2007).

Da kritisches Denken zu begründeten Urteilen und Handlungen führen soll, ist die Auseinandersetzung mit dem eigenen Selbst und den eigenen Vorstellungen und Handlungen unerlässlich (vgl. Hunter 2009, S. 3). Jahn (2013) schließt daraus, dass „Kritisches Denken und (Selbst-)Reflexion [...] nicht voneinander zu trennen [sind]“ (ebd., S. 162) und sieht das Reflektieren dabei als Bestandteil des kritischen Denkens. Auch nach Dubs (1992) beinhaltet kritisches Denken das Reflektieren über persönliche Annahmen, Werte und „über die Konsistenz von Denken und Verhalten“ (ebd., S. 43). Während für Moon (2007, S. 129) das Ziel des kritischen Denkens eher darin liegt, ein wohlüberlegtes Urteil zu treffen, das auf Beweisen beruht, werden beim Reflektieren verschiedene Ansätze untersucht, die Potenziale für die Urteilsbildung enthalten. Der Fokus dabei liegt weniger auf bestehende Tatsachen, sondern vielmehr auf unterschiedliche Wahrnehmungen. Moon fasst die Unterscheidung der beiden Begriffe wie folgt zusammen: „[R]eflection might therefore be seen as exploratory, whereas critical thinking is still a process that focuses on the making of a judgement, albeit with only surface issues described and under consideration“ (ebd., S. 129).

In der hier vorliegenden Studie wird sich an Moons Verständnis vom kritischen Denken orientiert. Das kritische Denken ist deshalb rationaler ausgelegt und beinhaltet weniger die individuellen Wahrnehmungen und Wertungen, die bei Reflektierprozessen – insbesondere bei der Selbstreflexion – von Bedeutung sind. Reflektieren steht folglich in Verbindung mit dem kritischen Denken, jedoch mit einer anderen Zielverfolgung. Kluge et al. bezeichnen das Reflektieren als „kritische Brille“ (o. J., S. 4) während der Analyse einer Situation, mit deren Hilfe

⁷ Dies entspricht dem Verständnis der *reflection-on-action* nach Schön (1983).

neue Denk- und Handlungsalternativen entwickelt werden. Das Reflektieren ermöglicht dabei die Berücksichtigung von Vorerfahrungen, Vorwissen und menschlicher Skepsis vor Neuem, die wichtig für kritische Überlegungen sind. Zusammengefasst ist das Reflektieren also Bestandteil des kritischen Denkens, das am Ende zu einer rationalen Entscheidung führt, die beim reinen Reflektieren nicht unbedingt erreicht wird. Vielmehr steht dabei die Untersuchung von Wissen und Handlungen aus unterschiedlichen Perspektiven und die Abwägung dieser im Fokus. Rationale Entscheidungen sind dabei nur schwer möglich, da individuelle Präferenzen und Werte die Ansichten beeinflussen, wie die Definition des Begriffs *Reflektieren*, der für diese Arbeit verwendet wird, zeigt: Reflektieren bedeutet Sachverhalte, Wissen, eigene und fremde Verhaltensweisen und Einstellungen kritisch zu hinterfragen und diese daraufhin zu überprüfen, warum sie ihre Gültigkeit und ihren Wert besitzen (oder auch nicht). Unter Reflektieren kann also ein tiefgreifendes Nachdenken über bestimmte Themen und Handlungen verstanden werden.

2.2 Bedeutung von Reflektierfähigkeiten im Geographieunterricht

Für den Erwerb von Kompetenzen ist die Fähigkeit Reflektieren zu können eine nicht zu vernachlässigende Größe (vgl. Hilzensauer 2008; Rychen 2003). Im Geographieunterricht ist die Reflektierfähigkeit eng verknüpft mit dem Erwerb spezieller Kompetenzen. Im Folgenden wird die eben angedeutete Wichtigkeit von Reflektieren und den Teilfertigkeiten der Reflexion und Reflexivität für den Geographieunterricht genauer erläutert. Dafür wird vorab auf den Kompetenzbegriff eingegangen und die Kompetenzen, die für das Fach Geographie definiert werden, vorgestellt.

Neben dem Wissenserwerb steht in der schulischen Bildung auch der Erwerb von Kompetenzen im Mittelpunkt. Um Schülerinnen und Schüler optimal auf die „funktionalen Anforderungen der Lebens- und Arbeitswelt“ (Klieme 2004, S. 11) vorzubereiten, wurden seit der Jahrtausendwende Bildungsstandards für einzelne Fächer definiert. Bildungsstandards können als eine „Form der Festlegung von Zielen für schulische Lehr- und Lernprozesse verstanden werden“ (ebd., S.10). Sie bestehen aus mehreren Kompetenzen, die wiederum abhängig vom jeweiligen Schulfach sind. Der Kompetenzbegriff wird im Bildungsbereich nach dem Verständnis nach Weinert verwendet und Kompetenzen folgendermaßen definiert:

Kompetenzen sind die bei Individuen verfügbaren oder durch sie erlernbaren kognitiven Fähigkeiten und Fertigkeiten, um bestimmte Probleme zu lösen, sowie die damit verbundenen motivationalen, volitionalen [...] und sozialen Bereitschaften und Fähigkeiten, um die Problemlösungen in variablen Situationen erfolgreich und verantwortungsvoll nutzen zu können. (Weinert 2002, S. 27 f.)

Im Laufe der Schullaufbahn soll Schülerinnen und Schülern ausreichend Raum und Zeit gegeben werden, Kompetenzen anwenden und weiterentwickeln zu können. Lehrpläne spielen

neben den Bildungsstandards weiterhin eine wichtige Rolle, da sie Wegbereiter für die Kompetenzförderung sind. War man vor der Einführung der Bildungsstandards hauptsächlich am Input interessiert, muss nun sowohl der Input- als auch der Output-Orientierung gleichermaßen Beachtung gegeben werden (vgl. DGfG 2014, S. 2). Für einige Fächer, jedoch nicht für das Schulfach Geographie, hat sich die Ständige Konferenz der Kultusminister der Länder der Bundesrepublik Deutschland der Kompetenzdefinierung und Festlegung der Bildungsstandards gewidmet. Um dennoch „[...] die Qualität des Bildungsprozesses im Schulfach Geographie zu sichern und weiter zu entwickeln sowie die Lehrpläne auf ein bundesweit einheitliches Fundament zu stellen [...]“ (ebd., S. 2), hat sich die Deutsche Gesellschaft für Geographie, bestehend aus Fachdidaktikern, Schulpraktikern und Wissenschaftlern des Faches Geographie, dieser Aufgabe angenommen. Grundlage für die sechs Kompetenzbereiche für das Fach Geographie sind Kliemes Expertise *Zur Entwicklung nationaler Bildungsstandards* (2007) und weitere geographische Dokumente wie die Internationale Charta oder das Curriculum 2000+. Da das Fach Geographie sowohl natur- als auch gesellschaftswissenschaftlich ausgelegt ist, werden in den Bildungsstandards auch Kompetenzen aus beiden Bereichen berücksichtigt. Neben den Kompetenzbereichen Fachwissen, Methoden, Kommunikation und Beurteilung/Bewertung, die auch in anderen naturwissenschaftlichen Fächern ausgewiesen werden, besteht das geographische Kompetenzgerüst auch noch aus den Kompetenzbereichen Handlung – auch vorhanden in rein gesellschaftlichen Fächern – und Räumliche Orientierung. Letzteres findet sich nur im Fach Geographie und stellt somit ein Charakteristikum des Faches dar (vgl. DGfG 2014, S. 8). Neben den Kompetenzbereichen werden drei Anforderungsbereiche in Anlehnung an die *Einheitlichen Anforderungen in der Abiturprüfung* (KMK 2005)⁸ vorgestellt, die der Lehrkraft als Orientierung bei Aufgabenformulierungen dienen sollen. Jede Kompetenz kann in allen drei Anforderungsbereichen gefordert und gefördert werden. Anforderungsbereich I umfasst den Bereich der Reproduktion, Anforderungsbereich II die Bereiche Reorganisation und Transfer und Anforderungsbereich III den Bereich Reflexion und Problemlösung (vgl. DGfG 2014, S. 31).

Für Hilzensauer (2008) und Rychen (2003) gilt das Reflektieren als Grundprinzip für die Ausbildung von Kompetenzen. Für die Bildungsstandards des Faches Geographie spielen Reflektiervorgänge – genauso wie in der Kompetenzentwicklung anderer Fächer – deshalb eine wichtige Rolle. Daneben gibt es weitere quantitative als auch qualitative Belege für die Wichtigkeit von Reflexionsvorgängen innerhalb der geographischen Bildungsstandards, die im Folgenden näher untersucht werden sollen.

⁸ KMK steht für Kultusministerkonferenz

Im ausformulierten Text der Deutschen Gesellschaft für Geographie (2014) wird 21 Mal Bezug aufs Reflektieren genommen.⁹ Der Begriff wird zwar häufig explizit genannt, allerdings wird an keiner Stelle geklärt, was man darunter in diesem Kontext eigentlich versteht. Hier wird die zuvor beschriebene Problematik der undefinierten Begriffsverwendung deutlich (vgl. Kapitel 2.1.1). Nimmt man Brendels (2017) Begriffsverständnisse von Reflexion und Reflexivität¹⁰ als Grundlage zur genaueren Analyse der Bildungsstandards, wird deutlich, dass beide Arten für den Kompetenzerwerb im Fach Geographie von Bedeutung sind. Besonders bei Aufgaben, die dem Anforderungsbereich III entsprechen, spielen Reflexion und Reflexivität eine wichtige Rolle:

Anforderungsbereich III (Reflexion und Problemlösung) umfasst den selbstständigen reflexiven Umgang mit neuen Problemstellungen, den eingesetzten Methoden sowie Verfahren und gewonnenen Erkenntnissen, um zu Begründungen, Deutungen, Folgerungen, Beurteilungen und Handlungsoptionen zu gelangen. Dies erfordert vor allem Leistungen der Reflexion und Problemlösung. (DGfG 2014, S. 31)

Damit Schülerinnen und Schüler mit der Komplexität und den kontrovers diskutierten, gesellschaftsrelevanten Themen unserer Zeit adäquat umgehen können, benötigen sie eine hohe Reflexions- und Reflexivitätsfähigkeit (vgl. Künzli et al. 2013). Denn erst, wenn sie ihre Überzeugungen und ihr Wissen zu unterschiedlichen Themen kritisch hinterfragen und unterschiedliche Perspektiven darauf erkennen und nachvollziehen können, sind sie in der Lage, begründet ihre Meinung mit anderen zu diskutieren. Damit sie im nächsten Schritt verantwortungsvoll handeln und die Gründe für ihr Verhalten unter Berücksichtigung moralischer, ethischer, nachhaltiger und globaler Aspekte analysieren können, müssen sie ihr Handeln stets metakognitiv untersuchen (vgl. ebd.). Um folglich dem Anforderungsbereich III gerecht zu werden, müssen Schülerinnen und Schüler stets sowohl Themen und ihr Wissen dazu als auch ihr eigenes Verhalten und das von anderen reflektieren.

Auch für den Erwerb einzelner Kompetenzen ist, wie bereits erwähnt, die Fähigkeit Reflektieren zu können unabdingbar (vgl. Hilzensauer 2008; Rychen 2003). Gerade für die in den Bil-

⁹ Folgende Begriffe werden explizit in den Bildungsstandards verwendet: „reflektiertes Heimatbewusstsein“ (DGfG 2014, S. 6), „reflektierte[r] Umgang mit Medien“ (ebd., S. 6), „reflektierte[] [...] Handlungsfähigkeit“ (ebd., S. 8), „Reflexion von Raumwahrnehmungen“ (ebd., S. 9 u. S. 18), „methodische[] Schritte [...] reflektieren“ (ebd., S. 21), „reflektieren [...] Entstehung und Entwicklungspotential einer Natur- oder Kulturlandschaft“ (ebd., S. 23), „reflektiert [...] zu einer Handlung entschließen“ (ebd., S. 26), „über Auswirkungen vollzogener oder geplanter Handlungen [...] reflektieren“ (ebd., S. 26), „Fähigkeit zur Reflexion der Handlungen“ (ebd., S. 28), „Reflexion und Problemlösung“ (ebd., S. 31), „reflexive[r] Umgang mit neuen Problemstellungen“ (ebd., S. 31), „Mensch-Umwelt-Beziehungen [...] reflektieren“ (ebd., S. 31), „Raumwahrnehmung und -konstruktion reflektieren“ (ebd., S. 31), „Erkenntniswege reflektieren“ (ebd., S. 31), „Wertmaßstäbe reflektieren“ (ebd., S. 31), „räuml. Auswirkungen mögl. Handlungen reflektieren“ (ebd., S. 31), „Reflexion der angewandten Wertmaßstäbe“ (ebd., S. 33), „Reflexion“ (ebd., S. 56), „reflektieren die Ergebnisse“ (ebd., S. 79).

¹⁰ Reflexion = das tiefgreifende Nachdenken über Inhalte; Reflexivität = das tiefgreifende Nachdenken über Handlungen

dungsstandards des Faches Geographie ausgewiesenen Kompetenzbereiche Beurteilung/Bewertung und Handlung ist das Reflektieren¹¹ ein Grundwerkzeug, ohne das die Kompetenzerweiterung nur schwer möglich ist. Damit Schülerinnen und Schüler „ausgewählte Situationen/Sachverhalte im Raum unter Anwendung geographischer/geowissenschaftlicher Kenntnisse [...] beurteilen [können]“ (DGfG 2014, S. 24) müssen sie zunächst Kriterien des Beurteilens kennen, wie beispielsweise die Multiperspektivität auf Sachverhalte oder konkrete Konzepte wie das Konzept des Nachhaltigkeitsdreiecks mit den Bereichen Soziales, Ökonomie und Ökologie. Außerdem müssen sie diese entsprechend anwenden und ihre daraufhin getroffenen Beurteilungen begründen können. Dies erfordert eine hohe Reflexionsfähigkeit über die Kriterien, den Sachverhalt selbst und die unterschiedlichen Ansichten und Meinungen darauf (vgl. Haan 2008). Wie später noch deutlich wird (vgl. Kapitel 2.4.2.1), gehört zur Reflexionsfähigkeit (gemeint ist hier das Nachdenken über das Thema selbst) auch die Fähigkeit des Beurteilens. Nach Brendel (2017), Chen et al. (2009) und Bain et al. (1999) müssen Schülerinnen und Schüler zum einen Sachverhalte und Entwicklungen vertieft begründen und zum anderen die Komplexität von Sachverhalten erkennen können. Letzteres ist erst durch das Erkennen und Nachvollziehen unterschiedlicher Perspektiven möglich, da es gerade durch diese Multiperspektivität zu keiner generell richtigen Ansicht kommen kann und somit vor allem gesellschaftsrelevante Themen als komplex und kontrovers diskutiert eingestuft werden (vgl. Haan 2008).

Auch die „Fähigkeit[en], ausgewählte geographisch/geowissenschaftlich relevante Informationen aus Medien kriteriengestützt zu beurteilen“ (DGfG 2014, S. 24) und „ausgewählte geographische/geowissenschaftliche Erkenntnisse und Sichtweisen hinsichtlich ihrer Bedeutung und Auswirkungen für die Gesellschaft angemessen zu beurteilen“ (ebd., S. 25), die im Kompetenzbereich Beurteilung/Bewertung als Standards ausgewiesen sind, setzen voraus, dass Schülerinnen und Schüler in der Lage sind, nachvollziehbare und wohl überlegte Urteile fällen zu *können*. Urteilen ist wiederum eine Fähigkeit, die u.a. durch das Reflektieren von Informationen aus unterschiedlichen medialen Quellen und der kriteriengeleiteten Reflexion über geographische Erkenntnisse besteht (vgl. Kölle 2017, S. 4). Auch Brendel (2017, S. 251) und Bain et al. (1999, S. 60) sehen das Urteilen als Teil der Reflexionsfähigkeit innerhalb ihrer Stufenmodelle. Der Zusammenhang zwischen Urteilen und Reflektieren betont die Notwendigkeit, Schülerinnen und Schülern mit den Grundlagen des Reflektierens vertraut zu machen, damit sie ihre Kompetenzen im Bereich Beurteilung/Bewertung entsprechend ausbauen können. Die Bewertung geographischer Sachverhalte unter Berücksichtigung bestimmter Werte, die ebenfalls als Bildungsstandard im Fach Geographie Erwähnung findet, ist ebenfalls nur möglich, wenn man den Weg zu seiner Wertungsentscheidung nachvollziehen und begründen

¹¹ Sowohl im Sinne von Reflexion als auch Reflexivität

kann. Dies setzt, wie bei allen Fähigkeiten innerhalb des Kompetenzbereichs Beurteilung/Bewertung, ein kritisches Hinterfragen der eigenen Überzeugungen voraus. Dies benötigt wiederum die Berücksichtigung unterschiedlicher Perspektiven und deren Ursachen (vgl. DGfG 2014, S. 23 ff.).

Zusammenfassend lässt sich für den Kompetenzbereich Beurteilung/Bewertung festhalten, dass Reflektieren eine Grundfähigkeit ist, die im direkten Zusammenhang mit der Ausbildung der einzelnen Teilfähigkeiten innerhalb dieses Bereiches steht. Erst wenn kritisch über Sachverhalte, Themen, Überzeugungen und Meinungen reflektiert wird¹², können nachvollziehbare Urteile und Wertungen gebildet werden.

Auch im Kompetenzbereich Handlung der Bildungsstandards des Faches Geographie spielen Reflexionsprozesse eine elementare Rolle. So lautet es darin explizit, dass „Schülerinnen und Schüler [...] im Unterricht nicht manipuliert oder zum Handeln genötigt werden [dürfen], sondern [...] sich reflektiert und begründet zu einer Handlung entschließen [sollen]“ (DGfG 2014, S. 26). Weiterhin wird vom Geographieunterricht gefordert, dass er „dazu befähigen und anregen [soll], über Auswirkungen vollzogener oder geplanter Handlungen und mögliche Alternativen zu reflektieren [...]“ (ebd.). Das Nachdenken über Handlungen und den Gründen dafür ist wichtig, um sein Verhalten für andere nachvollziehbar zu machen und es unter Heranziehen bestimmter Entscheidungskriterien gegenüber anderen rechtfertigen zu können.

Während in den Bildungsstandards für den Kompetenzbereich Beurteilung/Bewertung die Fähigkeit der Reflexion im Vordergrund steht, ist für den Kompetenzbereich Handlung die Fähigkeit der Reflexivität von großer Bedeutung. Deshalb ist es in Hinblick auf die Bildungsstandards für das Fach Geographie wichtig, sowohl Reflexion als auch Reflexivität, d.h. sowohl das tiefgreifende Nachdenken über Themen als auch über das eigene Verhalten bei Schülerinnen und Schülern für die Förderung einer ganzheitlichen Reflektierfähigkeit zu beachten. Doch auch für den Erwerb weiterer Kompetenzen ist die Fähigkeit des Reflektierens unabdingbar. Denn um den Leitziele des Faches Geographie gerecht zu werden, indem eine „reflektierte[], ethisch begründete[] und verantwortungsbewusste[] raumbezogene Handlungsfähigkeit“ (DGfG 2014, S. 8) ausgebildet wird, bedarf es, wie der Wortlaut bereits festlegt, die Fähigkeit sowohl über Themen als auch über Verhalten reflektieren zu können. Insgesamt betrachtet kann Reflektieren hinsichtlich der Bildungsstandards für das Fach Geographie als eine Grundfähigkeit gesehen werden, die Voraussetzung für den Erwerb spezieller Kompetenzen ist.

¹² Dies entspricht dem Verständnis der Reflexion nach Brendel (2017).

2.3 Bedeutung von Reflektierfähigkeiten im Bereich BNE und dem Globalen Lernen

Die Fähigkeit des Reflektierens ist nicht nur in den Bildungsstandards für das Fach Geographie relevant, auch im Rahmen einer Bildung für nachhaltige Entwicklung (BNE) und beim Globalen Lernen sind Reflektierprozesse unerlässlich, um deren definierte Zielkompetenzen ausbilden zu können. Zunächst soll kurz allgemein auf die beiden Konzepte Bildung für nachhaltige Entwicklung und Globales Lernen eingegangen werden, bevor die Bedeutung des Reflektierens darin herausgearbeitet wird.

2.3.1 Bildung für nachhaltige Entwicklung und Globales Lernen

Die beiden Konzepte Bildung für nachhaltige Entwicklung und Globales Lernen sind eng miteinander verzahnt, ihre Beziehungen werden jedoch von Autorinnen und Autoren unterschiedlich aufgefasst (vgl. Scheunpflug und Asbrand 2006). Einige betrachten BNE als übergeordnetes Konzept und Globales Lernen als einen Teilbereich davon (vgl. Overwien 2009), andere wiederum sehen beide Konzepte als gleichgestellt innerhalb einer nachhaltigen Entwicklung (vgl. Nestvogel 2002). Im Rahmen dieser Studie werden die beiden Konzepte als gleichrangig gesehen mit unterschiedlichen Ursprüngen, jedoch ähnlichen Zielen. Während das Konzept BNE aus der Umweltbildung weiterentwickelt wurde, entstand das Globale Lernen aus einer entwicklungspolitischen Initiative heraus (vgl. Rieckmann 2010, S. 11). Ersteres hat seinen Ursprung bei der UNO-Konferenz im Jahr 1992 in Rio de Janeiro, bei der sich 180 Staaten in der Agenda 21 darauf einigten, Maßnahmen einzuführen, um dem Prinzip einer nachhaltigen Entwicklung gerecht zu werden (vgl. Bahr 2013, S. 17; Schockemöhle und Schrüfer 2015). In den folgenden Jahren wurde die Thematik vertieft aufgegriffen und weitere Beschlüsse und Maßnahmen umgesetzt, darunter auch speziell für den Bildungsbereich (vgl. Bahr 2013, S. 18). Globales Lernen kann hingegen nach Scheunpflug (2001) „als die pädagogische Reaktion auf die Entwicklungstatsache zur Weltgesellschaft“ (ebd., S. 87) gesehen werden. Fokus des Konzepts ist die Frage nach „weltweiter Gerechtigkeit und den wirtschaftlichen und sozialen Möglichkeiten eines Zusammenlebens auf diesem Planeten (vgl. ebd. S. 92). Beide „zielen [...] auf den Erwerb von Schlüsselkompetenzen ab, die Menschen zur nachhaltigen Gestaltung der (Welt-) Gesellschaft befähigen“ (Rieckmann 2010, S. 11) und „möchten [...] einen Beitrag zur Förderung einer nachhaltigen Entwicklung leisten“ (ebd.). Beide Konzepte finden im schulischen Kontext Einzug und werden als fächerübergreifende Bildungsaufgaben gesehen.

Im Folgenden wird aufgrund der gleichen Zielverfolgung beider Konzepte auf die Bedeutung von Reflektierfähigkeiten in beiden Bereichen eingegangen. Zunächst werden zwei Auffassungen vorgestellt, wie BNE zur nachhaltigen Gesellschaftsentwicklung beitragen kann und inwieweit dabei Reflektierprozesse involviert sind. Anschließend wird der Orientierungsrahmen für den Lernbereich Globale Entwicklung im Rahmen einer Bildung für nachhaltige Entwicklung aufgegriffen und die Reflektierfähigkeiten, die dabei von Bedeutung sind, näher erläutert.

2.3.2 Reflektierfähigkeiten innerhalb der BNE

Es besteht Konsens darüber, dass man bestimmte Schlüsselkompetenzen benötigt, um mit den ökologischen, ökonomischen und sozialen Problemen der heutigen Zeit adäquat umgehen zu können. Aufgrund der Globalisierung, des gesellschaftlichen Wandels und des technologischen Fortschritts benötigt jeder Fertigkeiten und Fähigkeiten, die es erlauben, in immer komplexer werdenden Situationen gerechtfertigte Entscheidungen fällen zu können (vgl. Rieckmann 2016, S. 89). BNE hat das Ziel, diese Schlüsselkompetenzen im Bildungskontext auszubilden und zu fördern (vgl. Rieckmann 2013). Innerhalb des letzten Jahrzehnts wurde intensiv den Fragen nachgegangen, welche Kompetenzen für eine BNE besonders wichtig sind und wie diese gefördert werden können. Welche Kompetenzen jedoch tatsächlich erworben werden und folglich welche Auswirkungen bestimmte Lehr-Lernprozesse auf den Erwerb von Kompetenzen haben, wurde und wird bisher nur kaum untersucht (vgl. Barth 2016; Barth und Rieckmann 2016).

Es existieren unterschiedliche Auffassungen, wie Bildung für nachhaltige Entwicklung dazu beitragen kann, eine nachhaltige Gesellschaftsentwicklung anzuregen. Vare und Scott (2007, S. 191) unterscheiden dabei zwei unterschiedliche Strömungen der Bildung für nachhaltige Entwicklung. Sie bezeichnen diese als *Education for Sustainable Development 1* und *Education for Sustainable Development 2*.

In der Strömung des *Education for Sustainable Development 1* wird angenommen, dass einer nachhaltigen Entwicklung eindeutig bestimmte Verhaltensweisen und Werte zugeordnet werden können. Ziel einer BNE hierbei ist es, diese identifizierten Werte zu vermitteln und entsprechendes Verhalten zu fördern (vgl. Vare und Scott 2007). Dieser Ansatz erinnert an die Umweltbildung, deren Ziel Wals (2011) folgendermaßen beschreibt: „Much environmental education aims at changing learner behaviour, including attitudes, beliefs and values“ (ebd., S. 178). Die Zielsetzung nach dem *Education for Sustainable Development 1* Ansatz findet sich auch in den Zielformulierungen der *UN-Dekade Bildung für nachhaltige Entwicklung* (2005-2014).

Education for Sustainable Development 2 möchte hingegen keine konkreten Werte und Verhaltensweisen vorgeben, sondern verfolgt einen reflexiveren Ansatz: Nach Vare und Scott (2007) sollen Situationen ermöglicht werden, in denen Individuen selbst über Probleme und Fragen bezüglich der Nachhaltigkeit nachdenken sollen und dadurch eigene Antworten und Überzeugungen finden. Grundlage für diesen Ansatz ist die Erkenntnis, dass bestimmte Verhaltensweisen nicht generell nachhaltiger sind als andere (vgl. Wals 2011).

In ESD2, we can't measure success in terms of environmental impacts because this is an open-ended process; outcomes will depend on people's unforeseen decisions in future, unforeseeable circumstances. But we can research the extent to which people have been informed and motivated, and been enabled to think critically and feel empowered to take responsibility. (Vare und Scott 2007, S. 194)

Unter dem Verständnis des *Education for Sustainable Development 2* Ansatzes steht die kritische Auseinandersetzung mit der Komplexität, Unsicherheit und den Widersprüchen, die Begleitelemente einer nachhaltigen Entwicklung sind. Nachhaltige Entwicklung wird hier als eigener Lernprozess verstanden (vgl. ebd.). Nach Mulder (2014) und Wals (2011) stehen der Aufbau von Kapazitäten und kritisches Denken als Fertigkeiten im Mittelpunkt, die benötigt werden um eigene Entscheidungen im Sinne einer nachhaltigen Entwicklung treffen zu können. Da Reflektieren und kritisches Denken verwandte Fähigkeiten sind (vgl. Kapitel 2.1.2), sollten auch Reflektierfähigkeiten im Kontext einer Bildung für nachhaltige Entwicklung gefördert werden, um dem Verständnis des *Education for Sustainable Development 2* Ansatzes gerecht zu werden. De Haan (2008, S. 117 ff.) orientiert sich bei seiner Argumentation bezüglich der BNE ebenfalls am zweiten Ansatz. Ziel ist es demnach, kein Verhalten zu erzwingen, sondern Möglichkeiten zu eröffnen, um Entscheidungen für eigenes Verhalten mit bestimmten Zielen oder Absichten nachvollziehen und begründen zu können. Für die Erreichung der Ziele der BNE formuliert de Haan im Rahmen des geförderten Programms Transfer 21 zwölf Teilkompetenzen einer übergreifenden Gestaltungskompetenz. Dazu zählen folgende Kompetenzen (ebd.):

1. Weltoffen und neue Perspektiven integrierend Wissen aufbauen
2. Vorausschauend denken und handeln
3. Interdisziplinär Erkenntnisse gewinnen
4. Risiken, Gefahren und Unsicherheiten erkennen und abwägen können
5. Gemeinsam mit anderen planen und handeln können
6. An Entscheidungsprozessen partizipieren können
7. Sich und andere motivieren können aktiv zu werden
8. Zielkonflikte bei der Reflexion über Handlungsstrategien berücksichtigen können
9. Die eigenen Leitbilder und die anderer reflektieren können
10. Selbstständig planen und handeln können
11. Empathie und Solidarität für Benachteiligte zeigen können.
12. Vorstellungen von Gerechtigkeit als Entscheidungs- und Handlungsgrundlagen nutzen können

Bei den Teilkompetenzen 8 und 9 wird explizit die Reflektierfähigkeit aufgenommen. Das Dokument gibt jedoch keinen Hinweis darauf, was unter den verwendeten Begriffen *Reflektieren*

und *Reflexion* verstanden wird. Der Kontext lässt darauf schließen, dass sowohl Reflexion als auch Reflexivität nach diesem Verständnis als Fertigkeiten erforderlich sind, um Gestaltungskompetenz ausbilden zu können. Neben der expliziten Nennung von Reflektieren spielen Reflektierfähigkeiten auch in den übrigen Teilkompetenzen eine Rolle. Denn zum Reflektieren gehören u.a. auch die Einnahme unterschiedlicher Perspektiven, das Abwägen unterschiedlicher Ansätze und der gerichtete Blick in die Zukunft (vgl. Kapitel 2.4.2).

2.3.3 Reflektierfähigkeit innerhalb des Orientierungsrahmens für den Lernbereich Globale Entwicklung

In Deutschland wurde nach dem Ausruf der Weltdekade das politische Programm Transfer21 (2004-2008) der Bund-Länder-Kommission für Bildungsplanung und Forschungsförderung ins Leben gerufen, das auf BLK 21 (1999-2004) aufbaut. Inzwischen sind BNE und Globales Lernen feste Bestandteile der schulischen Bildung und sollten als fächerübergreifende Bildungskonzepte realisiert werden. Der Orientierungsrahmen für den Lernbereich Globale Entwicklung im Rahmen einer Bildung für nachhaltige Entwicklung¹³ soll dabei helfen, dieser verbindlichen Aufgabe gerecht zu werden und BNE im Unterricht eines jeden Faches zu implementieren (vgl. KMK und BMZ 2015, S. 84)¹⁴. Grundlage für den Orientierungsrahmen und die darin definierten Kernkompetenzen ist das Konzept der Gestaltungskompetenz nach de Haan (2008).

Das übergeordnete Bildungsziel im Lernbereich Globale Entwicklung besteht darin, grundlegende Kompetenzen für eine zukunftsfähige Gestaltung des privaten und beruflichen Lebens, für die Mitwirkung in der Gesellschaft und die Mitverantwortung im globalen Rahmen zu erwerben. (KMK und BMZ 2015, S. 18).

Dieses Ziel soll durch den Erwerb von elf Kernkompetenzen, die sich den Bereichen *Erkennen*, *Bewerten* und *Handeln* zuordnen lassen, erreicht werden. Im Gegensatz zu den zwölf Teilgestaltungskompetenzen nach de Haan sind die elf Kernkompetenzen des Orientierungsrahmens für den Lernbereich Globale Entwicklung im Rahmen einer BNE weniger von allgemeiner, überfachlicher Natur, sondern eher speziell auf nachhaltige Entwicklung und Globalisierung fokussiert. Sie ermöglichen es, direkt am kompetenzorientierten Fachunterricht anzuknüpfen und Verbindungen zu den fachbezogenen Teilkompetenzen herzustellen (vgl. KMK und BMZ 2015, S. 85 ff.). Besonders im Kompetenzbereich *Bewerten* sind Reflexionsprozesse ein wichtiger Bestandteil, weshalb dieser im Folgenden näher vorgestellt wird.

¹³ „Nachhaltige Entwicklung wird dabei nicht mit Wachstum gleichgesetzt, sondern als zukunftsfähige Entwicklung von Lebensqualität für alle verstanden – unter Berücksichtigung ökologischer Rahmenbedingungen“ (BMZ und KMK 2017, S. 18). Das Konzept BNE orientiert sich an den „vier Entwicklungsdimensionen Wirtschaft – Soziales – Umwelt – Politik“ (ebd., S. 19) einer nachhaltigen Entwicklung.

¹⁴ BMZ steht für Bundesministerium für wirtschaftliche Zusammenarbeit und Entwicklung

Im Orientierungsrahmen wird der Bereich *Bewerten* mit folgendem Wortlaut beschrieben: „Im Kompetenzbereich **Bewerten** (H.i.O.) geht es um kritische Reflexion und das Erkennen und Abwägen unterschiedlicher Werte sowie um Identitätsentwicklung auf der Grundlage wertorientierter Betrachtung.“ (KMK und BMZ 2015, S. 91). Hier wird bereits explizit auf die Wichtigkeit des kritischen Reflektierens eingegangen, das ebenfalls als Voraussetzung für die Kernkompetenz 8 „Solidarität und Mitverantwortung“ (ebd., S. 95) des Kompetenzbereichs Handeln gesehen wird (vgl. ebd.). Im weiteren Verlauf heißt es, dass eigene und fremde Werte erkannt, hinterfragt und verbalisiert werden sollen, wozu multiperspektivische und empathische Fähigkeiten benötigt werden. Beim Perspektivwechsel können unterschiedlichste Sichtweisen, z.B. von Akteuren, Betroffenen, verschiedenen Staaten oder Personengruppen, eingenommen werden (vgl. ebd.). Dass die Fähigkeit des Perspektivwechsels Teil der Reflexionsfähigkeit ist, wird in Kapitel 2.4.2.1 noch genauer erläutert. An dieser Stelle sei nur auf die bestehende Verbindung zwischen Multiperspektivität und Reflexion hingewiesen. Damit Individuen gestaltungskompetent werden, müssen sie ihre eigenen und fremde Überzeugungen und Handlungen identifizieren und nachvollziehen können. Dazu gehören sowohl die Fähigkeit der Reflexion als auch die Fähigkeit der Reflexivität. Denn erst wenn Entwicklungs- und Globalisierungsthemen in Bezug auf Normen, Werte, politische Vereinbarungen und Leitbilder kritisch reflektiert werden und dadurch eine bewusste Identifikation ermöglicht wird, wird man dem Verständnis der Bewertungskompetenz im Lernbereich der Globalen Entwicklung gerecht (vgl. KMK und BMZ 2015, S. 92).

Die drei Teilkompetenzen des Bereichs *Bewerten* enthalten bereits zweimal den direkten Bezug zu Reflexionsprozessen (vgl. Tabelle 1). Unsere Überzeugungen basieren auf dem, was wir lernen und als richtig erachten. Daraus ergeben sich unsere Ansichten und Handlungen. Das Reflektieren über eigene und fremde Vorstellungen zu bestimmten Themen ist deshalb essentiell, um die eigenen Schritte bis hin zu einer gezeigten oder auch nicht gezeigten Handlung auf metakognitiver Ebene nachvollziehen zu können. Erst wenn man sich dem Weg zu seinen eigenen Wertorientierungen bewusst ist, kann man andere Überzeugungen verstehen und begründen (vgl. Fishbein und Ajzen 1975).

Tabelle 1 Kernkompetenzen des Bereichs Bewerten des Orientierungsrahmens für den Lernbereich Globale Entwicklung im Rahmen einer BNE (KMK und BMZ 2015, S. 96), eigene Überarbeitung

		Die Schülerinnen und Schüler können...
Kernkompetenz 5	Perspektivwechsel und Empathie	... sich eigene und fremde Wertorientierungen in ihrer Bedeutung für die Lebensgestaltung bewusst machen, würdigen und <u>reflektieren</u> .
Kernkompetenz 6	<u>Kritische Reflexion</u> und Stellungnahme	... durch <u>kritische Reflexion</u> zu Globalisierungs- und Entwicklungsfragen Stellung beziehen und sich dabei an der internationalen Konsensbildung, am Leitbild nachhaltiger Entwicklung und an den Menschenrechten orientieren.
Kernkompetenz 7	Beurteilen von Entwicklungsmaßnahmen	... Ansätze zur Beurteilung von Entwicklungsmaßnahmen (bei uns und in anderen Teilen der Welt) unter Berücksichtigung unterschiedlicher Interessen und Rahmenbedingungen erarbeiten und zu eigenständigen Bewertungen kommen.

Die Kernkompetenzen des Bereichs Bewerten des Orientierungsrahmens für den Lernbereich Globale Entwicklung im Rahmen einer BNE stehen im engen Zusammenhang zur Fähigkeit der Reflexion. Die Fähigkeit der Reflexivität ist in diesem Bereich nicht direkt nötig, dafür jedoch im Bereich Handeln, für deren Kompetenzentwicklung für „Solidarität und Mitverantwortung“ (KMK und BMZ 2015, S. 95) die Bewertungskompetenz als Voraussetzung gesehen wird. Handlungskompetenz bedeutet u. a. „die Bereitschaft, das eigene Verhalten mit den persönlichen Grundsätzen einer zukunftsfähigen Lebensgestaltung in Einklang zu bringen. Oft ist zwischen verschiedenen Handlungsweisen bewusst zu wählen, Normen- und Interessenskonflikte sind zu klären, die direkten und indirekten Folgen von Handlungen abzuschätzen“ (ebd., S. 92). Damit diese Aufgaben erfolgreich gelingen, ist die Fähigkeit der Reflexivität Voraussetzung (vgl. Kapitel 2.1.1.2).

Zusammenfassend lässt sich festhalten, dass innerhalb des Orientierungsrahmens für den Lernbereich Globale Entwicklung im Rahmen einer BNE im Bereich *Bewerten* die Fähigkeit der Reflexion explizit eine entscheidende Rolle spielt und im Bereich *Handeln* sowohl die Fähigkeit der Reflexion als auch der Reflexivität implizit Grundbestandteile davon sind. Für eine gelingende Implementierung von BNE im Unterricht müssen Schülerinnen und Schüler folglich in der Lage sein, reflektieren zu können. Doch nicht nur für die einzelnen Kompetenzbereiche innerhalb des Orientierungsrahmens lässt sich ein Bezug zwischen BNE bzw. Globalem Lernen und Reflektieren herstellen. Buddenberg (2014, S. 58) sieht Reflexivität¹⁵ neben Offenheit und Dynamik als zentralen Aspekt für eine BNE insgesamt, deren Kompetenzen vor allem zu einem selbstbestimmten Leben und zur Übernahme von Verantwortung in der Gesellschaft

¹⁵Sie selbst verwendet den Begriff *Reflexivität* mit dem gleichen Verständnis, wie er in dieser Arbeit Verwendung findet.

führen sollen. In Anlehnung an Michelsen ist dafür vor allem Reflexivität nötig: „Denn die Voraussetzung für einen sinnvollen Umgang mit einer sich stetig verändernden Welt ist die Fähigkeit, diese Veränderungen und sein eignes Handeln in Verbindung mit Erfahrungen reflektieren zu können (Michelsen 2008).“ (Buddeberg 2014, S. 59).

2.4 Messung von Reflektierfähigkeiten

Dieses Kapitel widmet sich den Voraussetzungen und Methoden zur Messung von Reflektierfähigkeiten. Zunächst wird auf Möglichkeiten zur Visualisierung von Reflektierprozessen eingegangen, bevor anschließend die Bedeutung von Kompetenzmodellen behandelt und konkrete Modelle zur Messung der beiden Teilbereiche des Reflektierens, der Reflexionsfähigkeit und der Reflexivitätsfähigkeit vorgestellt werden.

2.4.1 Schriftliche Erfassung von Reflektierfähigkeiten

Um Reflektierfähigkeiten messen zu können, müssen diese zunächst sichtbar gemacht werden. Denn das alleinige Beobachten von Handlungen reicht nicht aus, um ausreichende Einblicke in Denkprozesse zu erhalten (vgl. Copeland et al. 1993). Das Sichtbarmachen solcher Prozesse kann anhand von Texten oder Interviews ermöglicht werden, wie beispielsweise in den Arbeiten von Eysel (2006) oder Brendel (2017). Dabei ist die schriftliche Variante geeigneter, da Reflektieren eine Wiedergabe von Gedanken darstellt und durch das Schreiben eine bewusste Distanz zum Erlebten hergestellt werden kann und so Erfahrungen intensiv aufbereitet werden können (vgl. Hatton und Smith 1995, S. 49; Wyss 2013, S. 50 f.). Traditionelle schriftliche Zugänge sind beispielsweise das Lerntagebuch (vgl. Bräuer 2014), forschendes Lernen (vgl. Feindt 2007) und Portfolios (vgl. Jahncke et al. 2018). Daneben zählen mittlerweile Blogs zu einem weiteren, geeigneten Instrument, um schriftliche Räume für Reflektierprozesse zu ermöglichen (vgl. Gläser-Zikuda und Hascher 2007). Auf das Medium Blog wird in Kapitel 2.6 noch näher eingegangen. Neben der ersten Voraussetzung, dem Sichtbarmachen von Reflektierprozessen zur Messung von Reflektierfähigkeiten, müssen weitere Aspekte berücksichtigt werden, die nachfolgend ausführlich behandelt werden.

2.4.2 Modelle zur Bestimmung von Reflektierfähigkeiten

Da das Reflektieren als Fähigkeit verstanden wird und somit vom Grundprinzip einer Kompetenz entspricht, gelten für die Erfassung von Reflektierfähigkeiten die gleichen Bedingungen wie für die Messung von allgemeinen Kompetenzen. Die Messung von Kompetenzen stellt eine Herausforderung dar: „Die Verbindung von pädagogischen Konstrukten, psychologischen Kompetenzmodellen und Messverfahren ist eines der schwierigsten Probleme der Kompetenzforschung“ (Hartig und Klieme 2008, S. 24). Um Kompetenzen messen zu können, müssen sie zunächst modelliert werden, d.h. die zu messende Fähigkeit muss in ein Konstrukt mit

genauer Definition umgewandelt werden. Die Modellierung von Kompetenzen erfolgt in der Literatur durch unterschiedliche Ansätze. Nach Hartig und Klieme (2006, S. 132) kann man dabei zwischen Kompetenzstrukturmodellen und Kompetenzniveaumodellen unterscheiden. Ersteres teilt Kompetenzen in unterschiedliche Dimensionen ein, wodurch eine Hauptkompetenz durch das Vorhandensein verschiedener Teilkompetenzen beschrieben werden kann. Letzteres bezieht sich auf konkrete inhaltliche Beschreibungen einer Kompetenz und definiert unterschiedliche Niveaus dieser. Die Niveaus richten sich danach, welche Anforderungen Personen bezüglich der Kompetenz erfüllen (vgl. ebd.).

In Hinblick auf das Reflektieren existieren je nach Autor unterschiedliche Struktur- und Niveaumodelle. Wyss (2008, 6 f.) unterscheidet hinsichtlich der Dimensionen von Reflektieren zwischen zwei Ebenen und zwei Formen. Bei ersterem formuliert sie eine „geschlossene Reflexion“ (ebd.) und eine „offene Reflexion“ (ebd.). Die beiden Unterscheidungen leitet sie aus Mühlhausens (2006, S. 11) Vorschlag ab, zwischen einer Analyse und einer Reflexion zu unterscheiden. Unter einer geschlossenen Reflexion versteht sie die Integration von Materialien mit vordefinierten Kriterien, die den Reflexionsprozess steuern und folglich koordiniert ablaufen lassen. Eine offene Reflexion wird hingegen nicht durch Material gelenkt, sondern läuft frei ab und die Person bestimmt selbst, über was sie reflektiert (vgl. Wyss 2008, S. 6). Bei den Reflektierformen unterscheidet sie zwischen einer „Reflexion-in-der-Handlung“ (ebd.) und einer „Reflexion-über-die-Handlung“ (ebd.) basierend auf Schöns Konzept der reflexiven Praxis (1983). Beide Formen können bei den beiden vorgestellten Ebenen auftreten. Brendel (2017, S. 19) teilt das Reflektieren in die Dimensionen Reflexion und Reflexivität ein. Während sich die Reflexion auf „ein Nachdenken über Unterrichtsinhalte, Überzeugungen oder Vorstellungen“ (ebd.) bezieht, bezeichnet Reflexivität „ein Nachdenken über eigene Handlungen oder Handlungsalternativen“ (ebd.). Hinsichtlich der unterschiedlichen Stufen von Reflektierfähigkeiten gibt es ebenfalls verschiedene Niveaumodelle. Brendel (2017, 19 ff) gibt in ihrer Dissertationsarbeit einen Überblick über bestehende Modelle. Im Folgenden wird tabellarisch ein Überblick über Reflektierniveaumodelle und deren Einteilungen gegeben (vgl. Tabelle 2).

Tabelle 2 Niveaumodelle hinsichtlich Reflektierfähigkeiten

Niveaumodell	Eigenschaften
<i>Reflection Scale</i> nach Bain et al. (1999)	Fünfstufiges Niveaumodell mit den Niveaus <i>Reporting, Responding, Relating, Reasoning</i> und <i>Reconstructing</i>
<i>Reflective Judgement Model</i> nach King und Kitchenner (1994)	Siebenstufiges Niveaumodell mit den drei Hauptkategorien <i>Pre-reflective Thinking</i> (Stufe 1, 2 und 3), <i>Quasi Reflective Judgment Stages</i> (Stufe 4 und 5) und <i>Reflective Thinking</i> (Stufe 6 und 7)
<i>Reflection Levels</i> nach Chen et al. (2009)	Siebenstufiges Niveaumodell mit den Niveaus <i>Nonsense, Simple, Incomplete, Reporting, Responding, Relating, Reasoning</i> und <i>Reconstructing</i>
<i>Levels of Reflection</i> nach Hatton und Smith (1995)	Vierstufiges Niveaumodell mit den Niveaus <i>Descriptive Writing, Descriptive Reflection, Dialogic Reflection</i> und <i>Critical Reflection</i>
<i>Levels of Reflectivity</i> nach Mezirow (1981)	Siebenstufiges Niveaumodell mit den Niveaus <i>Reflectivity, Affective Reflectivity, Discriminant Reflectivity, Judgemental Reflectivity, Conceptual Reality, Psychic Reality, Theoretical Reflectivity</i>
<i>Levels of Reflective Thinking</i> Kember et al. (2000)	Vierstufiges Niveaumodell mit den Niveaus <i>habitual action/non-reflection, understanding, reflection</i> und <i>critical reflection</i>
<i>Levels of Reflectivity</i> nach Van Manen (1977)	Dreistufiges Niveaumodell mit den Niveaus <i>technical rationality, practical action</i> und <i>critical reflection</i>
Kompetenzstufen der Reflexivität nach Zimmermann und Welzel (2008)	Fünfstufiges Niveaumodell mit den Niveaus Sachbezogene Beschreibung, Handlungsbezogene Begründung, Analytische Abstraktion, Kritischer Diskurs, Professionalisierung
Stufenmodell reflexiven Denkens im Kontext Globalen Lernens nach Brendel (2017)	Fünfstufiges Niveaumodell mit den Niveaus Wiedergabe, Bezug nehmen, Zusammenhänge (aufgeteilt in Begründung und Vernetzung), Beurteilung und (Re)Konstruktion

Da LaBoskey (1993) vorschlägt, bei der Untersuchung von Reflektiervorgängen zwischen dem Fokus des Reflektierens und dem Level des Reflektierens zu unterscheiden, wird sich in dieser Studie sowohl an Struktur- als auch Niveaumodellen für die Modellierung der Reflektierfähigkeit orientiert. Die Reflektierfähigkeit setzt sich, wie sie in dieser Studie verstanden wird, aus geschlossenen Reflektiervorgängen im Sinne einer Reflexion-über-die-Handlung zusammen. Sie besitzt zudem die beiden Dimensionen Reflexion und Reflexivität in Anlehnung an Brendels Verständnis (2017). Die Reflektierfähigkeit wird folglich mithilfe dreier Eigenschaften in ein Strukturmodell eingeordnet (vgl. Tabelle 3).

Tabelle 3 Strukturmodell der Reflektierfähigkeit

Reflektierfähigkeit	
Reflexion	Reflexivität
Reflektieren über Inhalte und Wissen	Reflektieren über Handlungen
Grundlage: Dewey (1933) Konzept	Grundlage: Schöns Konzept (1983)

Die Dimension Reflexionsfähigkeit der Reflektierfähigkeit wird im Rahmen dieser Studie in Anlehnung an das Niveaumodell von Brendel (2017) in fünf Niveaustufen eingeteilt. Auch die Dimension Reflexivitätsfähigkeit unterteilt sich in Anlehnung an das Niveaumodell von Zimmermann und Welzel (2008) in fünf Stufen. In den beiden folgenden Kapiteln wird noch genauer auf beide Modelle eingegangen. Neben einer detaillierten Beschreibung der Niveaustufen, werden die Entwicklungen beider Modelle und die Gründe für ihre Verwendung in dieser Studie aufgeführt.

2.4.2.1 Brendels Stufenmodell (2017) zur Bestimmung von Reflexionsfähigkeit

Brendel (2017, S. 251) entwickelt im Rahmen ihrer Dissertationsstudie ein Stufenmodell zum reflexiven Denken im Kontext Globalen Lernens, das Lehrkräften dazu dienen soll, die Reflexionsfähigkeiten ihrer Schülerinnen und Schüler einfach bestimmen zu können. Das Modell basiert auf dem ursprünglichen Stufenmodell von Bain et al. (1999) und der Überarbeitung dieses Modells von Chen et al. (2009). Nachfolgend sollen zunächst die beiden Ausgangsmodelle vorgestellt werden, bevor anschließend genauer auf die Adaptionen und Brendels Stufenmodell selbst eingegangen wird.

Das Stufenmodell von Bain et al. (1999) wurde im Rahmen einer Studie entwickelt, um die Reflexionsfähigkeiten von Lehramtsstudierenden während ihres Praktikums bestimmen zu können.¹⁶ Als Grundlage für die Analyse der Reflexionsfähigkeiten dienen die Überlegungen von LaBoskey (1993) und Tann (1993) zur Berücksichtigung zweier Dimensionen der Reflexion. Sie unterscheiden zwischen dem Fokus und dem Level der Reflexion. Zunächst muss bestimmt werden, welche Fokusse in einem Material zu finden sind, anschließend werden die auftretenden Levels der Reflexion für jeden Fokus bestimmt. Bain et al. verwenden eine Fünf-Punkte-Skala für die Einordnung der Reflexionsfähigkeiten in Reflexionslevels, deren Kategorien sie aus Literaturrecherchen entwickeln (vgl. Tabelle 4). Das Stufenmodell zeigt eine Orientierung an Blooms Taxonomie von Lernzielen (1956) und reicht von Beschreibungen bis zu komplexen Analysen innerhalb möglicher Reflexionsprozesse. Im Rahmen einer Studie von Chen et al. (2009) wurde das Modell als Grundlage zur Bestimmung der Reflexionsfähigkeiten von Biologiestudierenden verwendet. Die Autoren reduzieren die ursprünglichen Kriterien des

¹⁶ Die Studie wird im Rahmen des Forschungsstands noch ausführlicher dargestellt (vgl. Kapitel 2.7.2.2).

Modells deutlich und ergänzen es um drei weitere Stufen, die vor Level 1 angeordnet werden. Diese Stufen lauten *Nonsense*, *Simple* und *Incomplete* und werden eingeführt, da die Autoren der Meinung sind, dass es Aussagen gibt, die noch unter der ersten Stufe nach Bain et al. (1999) liegen (vgl. Tabelle 4). Jeder Stufe des Stufenmodells nach Chen et al. (2009) wird ein Scorewert zugeordnet, angefangen von 1 für die Stufe *Nonsense* bis 8 für die Stufe *Reconstructing*.

Tabelle 4 Übersicht der Reflexionsmodelle, die als Grundlage für Brendels Stufenmodell (2017) zur Bestimmung von Reflexionsfähigkeit dienen

Reflection levels	Five-point level of reflection scale by Bain et al. (1999)	Evaluation criteria by Chen et al. (2009)
Nonsense	- <i>Nicht vorhanden</i> -	- Reflection content no sense or missing
Simple	- <i>Nicht vorhanden</i> -	- Reflection content just like yes or no
Incomplete	- <i>Nicht vorhanden</i> -	- Reflection content incomplete
Reporting	- The student describes, reports or re-tells with minimal transformation, no added observations or insights	- Only repeated the content already in article
Responding	<ul style="list-style-type: none"> - The student uses the source data in some way, but with little transformation or conceptualisation - The student makes an observation or judgment without making any further inferences or detailing the reasons for the judgement - The student asks a 'rethorical' question without attempting to answer it or consider alternatives - The student reports a feeling such as relief, anxiety, happiness, etc. 	<ul style="list-style-type: none"> - Just used few concepts - Stated observed phenomena, no reasoning - Stated personal affection only
Relating	<ul style="list-style-type: none"> - The student identifies aspects of the data which have personal meaning or which connect with their prior or current experience - The student seeks a superficial understanding of relationships - The student identifies something they are good at, something that they need to improve, a mistake they have made, or an area in which they have learned from their practical experience - The student gives a superficial explanation of the reason why something has happened or identifies something they need or plan to do or change 	<ul style="list-style-type: none"> - Stated relationships within the context - Explained the cause

Reasoning	<ul style="list-style-type: none"> - The student integrates the data into an appropriate relationship, e.g. with theoretical concepts, personal experience, involving a high level of transformation and conceptualisation - The student seeks a deep understanding of why something has happened - The student explores or analyses a concept, event or experience, asks questions and looks for answers, considers alternatives, speculates or hypothesises about why something is happening - The student attempts to explain their own or others' behaviour or feelings using their own insight, inferences or previous learning, with some depth of understanding 	<ul style="list-style-type: none"> - Detailed explanation of the rationale - Combined theory and practice
Reconstructing	<ul style="list-style-type: none"> - The student displays a high level of abstract thinking to generalise and/ or apply learning - The student draws an original conclusion from their reflections, generalises from their experience, extracts general principles, formulates a personal theory of teaching or takes a position on an issue - The student extracts and internalises the personal significance of their learning and/ or plans their own further learning on the basis of their reflections 	<ul style="list-style-type: none"> - Expressed high order of reconstructing Organized theory, rule and experience in a systematic approach

Das Stufenmodell reflexiven Denkens im Kontext Globalen Lernens nach Brendel (2017), das in der hier vorliegenden Studie Anwendung findet, orientiert sich einerseits am Stufenmodell nach Bain et al. (1999) und der Überarbeitung von Chen et al. (2009), andererseits an den Erkenntnissen aus Brendels eigener Studie. Brendels Modell besteht aus insgesamt fünf Stufen, wobei Stufe 3 in zwei Unterstufen aufgeteilt ist (vgl. Tabelle 5). Stufe 1 entspricht der „Wiedergabe“ (ebd., S. 251) und die höchste Stufe der „(Re)Konstruktion“ (ebd.). Auch hier entspricht die Anordnung der Stufen den Taxonomiestufen nach Bloom (1956), d.h. die Reflexionsfähigkeit wird nach oben hin höher, da tiefgründigere Reflexionsprozesse stattfinden. Hinsichtlich der ursprünglichen Stufenmodelle, auf denen Brendels Modell aufbaut, wurden folgende Modifizierungen vorgenommen (vgl. Brendel 2017, S. 249 ff.): Kriterien mit Reflexivitätsanteilen wurden vollständig entfernt, weitere Kriterien wurden aufgenommen und existierende Kriterien umformuliert und konkretisiert. Während im ursprünglichen Modell die Konzeptualisierung auf einem erhöhten bzw. hohen Abstraktionsniveau jeweils eine der Kriterien in den Stufen *Reasoning* und *Reconstructing* darstellt, wird dies bei Brendel als Grundvoraussetzung für das Erreichen der jeweiligen Stufen gesehen. In Stufe 3 nimmt Brendel das Krite-

rium Urteil mit simpler Begründung auf, da für sie die Einordnung einer einfachen Urteilsbegründung im ursprünglichen Modell unklar bleibt. Ergänzungen in den Stufen 4 und 5 entwickelt sie induktiv aus der Kodierung im Rahmen ihrer Studie zur Bestimmung des reflexiven Denkens von Schülerinnen und Schülern beim Bloggen. Weiterhin wird Stufe 3 in zwei Unterstufen eingeteilt, da sich bei der Kodierung des Materials in Brendels Studie zu große Unterschiede innerhalb dieser Stufe ergeben und sie deshalb weiter ausdifferenziert werden muss (vgl. ebd.).

Tabelle 5 Stufenmodell reflexiven Denkens im Kontext Globalen Lernens nach Brendel (2017, S. 251)

Reflexionsstufe		Kriterien zur Bestimmung der Reflexionsstufe
Stufe 1: Wiedergabe		Reine Reproduktion der Unterrichtsinhalte
Stufe 2: Bezug nehmen		Die Schülerin oder der Schüler: <ul style="list-style-type: none"> - transformiert oder konzeptualisiert die Unterrichtsinhalte geringfügig - äußert eine Beobachtung oder Meinung, jedoch ohne Begründung oder Folgerung - stellt rhetorische Fragen ohne Antwort - drückt Gefühle aus - nennt Lernzuwachs
Stufe 3: Zusammenhänge	Stufe 3a: Begründung	Die Schülerin oder der Schüler: <ul style="list-style-type: none"> - begründet Prozesse oder Sachverhalte oberflächlich, Handlungsbedarf wird erkannt - Urteil mit simpler Begründung
	Stufe 3b: Vernetzung	Die Schülerin oder der Schüler: <ul style="list-style-type: none"> - stellt persönlichen Zusammenhang zu den Inhalten her, z.B. durch Verbindung mit Vorwissen oder eigenen Erfahrungen - versucht oberflächlich Wechselbeziehungen darzustellen
Stufe 4: Beurteilung		<i>Voraussetzung: Konzeptualisierung auf gehobenem Niveau</i> Die Schülerin oder der Schüler: <ul style="list-style-type: none"> - verknüpft Inhalte mit z.B. Theorien oder persönlichen Erfahrungen - führt vertiefte Begründungen für Prozesse, Sachverhalte oder eigene und fremde Handlungen an - analysiert ein Problem, sucht nach Antworten und Alternativen, spekuliert oder stellt Hypothesen auf - untersucht den Zusammenhang zwischen Theorie und Praxis - erkennt die Komplexität des Sachverhalts

<p style="text-align: center;">Stufe 5: (Re)Konstruktion</p>	<p><i>Voraussetzung: Konzeptualisierung auf hohem Abstraktionsniveau</i></p> <p>Die Schülerin oder der Schüler:</p> <ul style="list-style-type: none"> - kommt zu persönlichem Fazit oder bezieht systematisch und differenziert Stellung - formuliert eine persönliche Theorie oder leitet Generalisierungen ab - erweitert und modifiziert die Unterrichtsergebnisse deutlich über das angestrebte Niveau hinaus
---	---

Das Stufenmodell nach Brendel (2017) wird für die Messung der Reflexionsfähigkeiten im Rahmen dieser Studie aus mehreren Gründen verwendet. Zum einen wurde das Modell speziell für die Messung von Reflexionsfähigkeiten im Kontext Globalen Lernens im Unterricht entwickelt. Zum anderen basiert es auf dem Modell nach Bain et al. (1999), das bereits häufig als Grundlage für die Entwicklung weiterer Niveaumodelle zur Messung von Reflektierfähigkeiten diente und deshalb als verlässliches Instrument gesehen werden kann (vgl. Carrington und Selva 2010; Chen et al. 2009; Henderson et al. 2004). Weiterhin stellt Brendel einen Bezug zu bestehenden Theorien des Reflektierens bei der Entwicklung ihres eigenen Niveaumodells her und ordnet ihres explizit dem Teilbereich der Reflexion zu und grenzt es von weiteren bestehenden Niveaumodellen ab. Sie definiert außerdem die Kriterien klar und nennt Ankerbeispiele, sodass für die eigene Kodierung eine transparente Orientierung gegeben wird (vgl. Brendel 2017, 251 ff.).

2.4.2.2 Stufenmodell nach Zimmermann und Welzel (2008) zur Bestimmung von Reflexivitätsfähigkeit

Das Niveaumodell, das als Ausgangspunkt für die Messung der Reflexivitätsfähigkeiten in dieser Studie verwendet wird, ist das Stufenmodell nach Zimmermann und Welzel (2008). Nachfolgend werden die einzelnen Stufen und die Grundlagen des Modells genauer vorgestellt. Zimmermann und Welzel (2008) verwenden ihre Stufeneinteilung der Reflexivitätsfähigkeit, um die Reflexivitätsleistungen von Erzieherinnen und Erziehern hinsichtlich einer naturwissenschaftlichen Frühförderung zu untersuchen. Sie differenzieren dabei fünf Stufen von Reflexivität. Stufe 1, die „sachbezogene Beschreibung“ (Zimmermann und Welzel 2008, S. 34) entspricht dabei der Stufe mit der geringsten Reflexivität und Stufe 5, die „Professionalisierung“ (ebd.) der mit der am höchsten erreichbaren Reflexivität (vgl. Tabelle 6).

Tabelle 6 Stufenmodell nach Zimmermann und Welzel (2008, S. 34) zur Bestimmung von Reflexivitätsfähigkeit

Stufe	Kriterien zur Bestimmung der Stufe
Stufe 1: Sachbezogene Beschreibung	<ul style="list-style-type: none"> - bezogen auf konkretes pädagogisches Angebot - einfache Wiedergabe von Handlungsabläufen - Beschreibung von Situationen ohne zusätzliche Kommentare, ohne eigene Gedanken - Erste Distanzierung vom konkreten Tun ohne Begründung, keine explizite Reflexion
Stufe 2: Handlungsbezogene Begründung	<ul style="list-style-type: none"> - Beschreibung von Sachverhalten - Sinngebung mit Begründungen oder Handlungsalternativen werden aufgeführt - Konkrete Handlungsebene - Innere Distanzierung vom bloßen Handeln, da Formulierung der eigenen Perspektive vorhanden
Stufe 3: Analytische Abstraktion	<ul style="list-style-type: none"> - Gemachte Erfahrungen werden unter Berücksichtigung mehrerer Perspektiven (z.B. aus der Sicht von Erzieherin und Kind; Vergangenheit und Gegenwart in Bezug auf eigenes pädagogisches Handeln) analysiert - Elemente und Strukturen werden auf einer allgemeineren Ebene erkannt, identifiziert und zueinander in Beziehung gesetzt - aus dieser Analyse heraus werden Konsequenzen im Hinblick auf zukünftiges Verhalten, zukünftige Aktivitäten gezogen - ein neues, erweitertes Verständnis wird entwickelt
Stufe 4: Kritischer Diskurs	<ul style="list-style-type: none"> - die Erkenntnisse werden im Hinblick auf über geordnete Werte, Prinzipien und/oder auf relevante Theorien hin bewertet - es erfolgt eine kritische Auseinandersetzung damit - Zunahme der Fähigkeit zum Erkennen eigener - Anteile und Deutungen in Interaktionen
Stufe 5: Professionalisierung	<ul style="list-style-type: none"> - Erkennen des eigenen Urteils- Handlungs-Hiatus UND Durchdenken, Planen von zielorientierten Verbesserungsstrategien - konkrete Verbesserungsmaßnahmen (am eigenen Tun, an professionellem Umfeld) werden initiiert - festgefahrene Rollenverständnisse werden thematisiert und konstruktiv in Frage gestellt, andere Handlungsmöglichkeiten diesbezüglich ausgelotet UND ausprobiert - Herausdestillieren konstruktiver, unterstützender Fortbildungselemente für NFFK

Grundlage des Modells sind die Reflektierarten nach Hatton und Smith (1995). Die beiden Autoren entwickelten ihr Modell im Rahmen einer Untersuchung des Gegenstandes Reflektieren innerhalb der Literatur und der Analyse schriftlicher Berichte von Lehramtsstudierenden

(vgl. Kapitel 2.7.2.2). Dabei identifizieren die Autoren vier unterschiedliche Arten von Reflexionsprozessen, die wie folgt lauten:

Tabelle 7 Reflektierarten nach Hatton und Smith (1995, S. 48 f.)

Stufe	Kriterien zur Bestimmung der Stufe
Descriptive Writing	<ul style="list-style-type: none"> - Not reflective. - Description of events that occurred/report of literature. - No attempt to provide reasons/justification for events.
Descriptive Reflection	<ul style="list-style-type: none"> - Reflective, not only a description of events but some attempt to provide reason justification for events or actions but in a reportive or descriptive way. For example, "I chose this problem-solving activity because I believe that students should be active rather than passive learners." - Recognition of alternate viewpoints in the research and literature which are reported [...]. - Two forms: Reflection based generally on one perspective/factor as rationale. Reflection is based on the recognition of multiple factors and perspectives.
Dialogic Reflection	<ul style="list-style-type: none"> - Demonstrates a "stepping back" from the events/actions leading to a different level of mulling about, discourse with self and exploring the experience, events, and actions using qualities of judgements and possible alternatives for explaining and hypothesising. Such reflection is analytical or/and integrative of factors and perspectives and may recognise inconsistencies in attempting to provide rationales and critique, for example. "While I had planned to use mainly written text materials I became aware very quickly that a number of students did not respond to these. Thinking about this now there may have been several reasons for this. A number of students did not respond to these. Thinking about this now there may have been several reasons for this. A number of students, while reasonably proficient in English, even though they had been NESB learners, may still have lacked some confidence in handling the level of language in the text. Alternatively, a number of students may have been visual and tactile learners. In any case I found that I had to employ more concrete activities in my teaching." - Two forms, as in (a) and (b) above.
Critical Reflection	<ul style="list-style-type: none"> - Demonstrates an awareness that actions and events are not only located in, and explicable by, reference to multiple perspectives but are located in, and influenced by multiple historical, and socio-political contexts. For example, 'What must be recognised, however, is that the issues of student management experienced with this class can only be understood within the wider structural locations of power relationships established between teachers and students in schools as social institution based upon the principle of control'.

Die vier identifizierten Arten des Reflektierens sind weniger als Stufen, sondern vielmehr als unterschiedliche Formen ohne Wertungen zu sehen. Die Autoren erkennen jedoch Entwicklungsmöglichkeiten für konkrete Stufenmodelle, um mit einem solchen Modell mehr Aussagen über die Reflexivitätsfähigkeiten treffen zu können (vgl. Hatton und Smith 1995, S. 45 f.). Der Empfehlung folgend wurden die vier Reflexivitätsarten nach Hatton und Smith bereits in vielen Forschungsarbeiten als Grundlage zur Identifizierung unterschiedlicher Reflexivitätsniveaus verwendet und zum Teil nur leicht modifiziert (vgl. Abels 2011; Brendel 2017; Eysel 2006;

Korthagen 2002; Zimmermann und Welzel 2008). Eysel (2006) untersucht Portfolios hinsichtlich gezeigter Reflexivitätsfähigkeiten, um die Wirkung eines Veranstaltungskonzepts im Rahmen eines fächerübergreifenden Projektseminars messen zu können. Hierbei verwendet sie die ursprünglichen Reflexivitätsarten von Hatton und Smith (1995) und ordnet sie einem Stufenmodell zu. Die Entwicklung des Stufenmodells begründet sie dabei auf den Erkenntnissen von Korthagen (2002), der einen Überblick über den Bereich des Reflektierens in der Literatur gibt und sich ebenfalls an Hatton und Smith (1995) orientiert.

Insgesamt kann festgehalten werden, dass das Urmodell von Hatton und Smith (1995) weiterhin präsent in der Reflexivitätsforschung ist und die Einteilung der Stufen von Eysel (2006) nach den Worten von Rahm und Lunkenbein (2014) „derzeit als richtungsweisend gelten [kann]“ (ebd., S. 246). Das Modell der Reflexivitätsstufen nach Zimmermann und Welzel (2008) wird als Grundlage für die Bestimmung der Reflexivitätsfähigkeiten im Rahmen dieser Studie gewählt, da dem Modell eine ausführliche Beschreibung und Kodierhilfen zu Grunde liegen, an denen sich für die Auswertung in dieser Studie orientiert werden kann. Außerdem kann somit untersucht werden, ob Schülerinnen und Schüler sich hinsichtlich nachhaltigen Handelns bereits professionalisieren können oder ob die höchste Stufe erst zu einem späteren Zeitpunkt erreicht wird. Die Stufen nach Eysel (2006) wurden bei dem Modell von Zimmermann und Welzel (2008) nur leicht modifiziert, sodass dieses Modell ebenfalls als reliabel gesehen werden kann. Es gilt zu beachten, dass die beschriebenen Reflexivitätsstufen bisher nur für die Bestimmung von Reflexivitätsfähigkeiten im beruflich pädagogischen Kontext, wie dem Lehrberuf (vgl. Brendel 2017) oder dem Erzieherberuf (vgl. Zimmermann und Welzel 2008) eingesetzt werden. In der hier vorliegenden Studie wird das Modell nach Zimmermann und Welzel (2008) verwendet, um Reflexivitätsfähigkeiten von Schülerinnen und Schülern bezüglich ihres Handelns zu unterschiedlichen Themenfeldern zu untersuchen. Im Kontext von außerberuflichen, alltäglichen Verhaltensgewohnheiten, wie beispielsweise der regelmäßige Einkauf oder die Nutzung von Verkehrsmitteln, wurden das Modell von Hatton und Smith (1995) und dessen Adaptionen nach aktuellem Wissensstand der Forscherin noch nicht zum Auswerten von Reflexivitätsniveaus verwendet. Für diesen Bereich kann es jedoch auch als sinnvolles Instrument eingesetzt werden. Denn nicht nur berufliches Handeln muss professionalisiert werden, sondern auch Routinen oder konkrete Alltagshandlungen sind es wert, kritisch hinterfragt zu werden, damit festgefahrene Verständnisse neu überdacht und gegebenenfalls modifiziert werden. Gerade in Bezug auf Nachhaltigkeit ist es wichtig, kritisches Denken und Reflektieren zu fördern, damit künftige Handlungsentscheidungen wohl überlegt und begründet getroffen werden können und Verhaltensgewohnheiten nicht als Normalität anzusehen (vgl. Vare und Scott 2007, S. 194). Das Hinterfragen von Handlungen und dem eigenen Verhalten ist logischerweise nicht bei allen Gewohnheiten erforderlich, da manche Gewohnheiten, wie beispielsweise das Blinken vor dem Abbiegen mit dem Auto oder das Absperren der Tür beim

Verlassen der Wohnung, praktisch sind und geistige Energie sparen. Doch „es lohnt sich [] in vielen Fällen, Gewohnheiten und Einstellungen zu hinterfragen“ (Kopp 2015, S. 24) und sie nicht als normal anzusehen, nur weil es schon immer so gemacht wurde. Alternativhandlungen und Innovationen benötigen zunächst ein kritisches Prüfen bestehender Verhaltensweisen mit der sich immer wieder stellenden Frage, ob „das die bestmögliche Variante zu handeln [sei]?“ (ebd., S. 23). Um das Bewusstsein für nachhaltiges Handeln zu fördern, ist es deshalb nötig, sich Gedanken über sein Verhalten diesbezüglich zu machen und die Gründe dafür nachvollziehen zu können. Damit dies erfolgreich geschehen kann und im Idealfall professionalisiert wird, ist eine hohe Reflexivitätsfähigkeit wichtig – ähnlich wie in beruflichen Handlungssituationen. Folglich können die Reflexivitätsstufen nach Hatton und Smith (1995) mit leichten Modifizierungen (vgl. Kapitel 3.4.1) auch im Kontext einer BNE zur Auswertung der Reflexivitätsfähigkeit verwendet werden.

2.5 Förderung von Reflektierfähigkeiten

Reflektieren ist eine Fähigkeit, die sowohl für Lehrende als auch Lernende unabdingbar ist, um den Anforderungen der heutigen Zeit gerecht zu werden. Denn Reflektieren ist Bestandteil des selbstorganisierten Lernens, das Voraussetzung ist, um mit der Komplexität und den sich ständig verändernden Situationen umgehen zu können (vgl. Jenert 2008, S. 1). Somit ist es ratsam, die Fähigkeit des Reflektierens während der schulischen Ausbildung bereits von Beginn an zu fördern und es auch in spätere Ausbildungsprogramme – vor allem auch in der Ausbildung von Lehrerinnen und Lehrern (vgl. Felten 2005) – zu integrieren. Im Folgenden wird auf Theorien zur Förderung von Reflektierfähigkeiten eingegangen und anschließend die beiden Methoden genauer vorgestellt, die im Rahmen der hier vorliegenden Studie als zu untersuchende Reflektierhilfen gewählt wurden.

2.5.1 Theorien zur Förderung von Reflexionsprozessen

Mann et al. (2007, S. 614) sind überzeugt, dass Reflektierfähigkeiten gelehrt und gelernt werden können. Damit dies erfolgreich gelingen kann, müssen den Lernenden ausreichend Zeit zum Reflektieren gegeben und eine Atmosphäre geschaffen werden, die Reflektieren überhaupt erst ermöglicht. Zusätzlich dazu ist eine professionelle Unterstützung bei den Reflektierprozessen zu empfehlen. Um Reflektierprozesse sichtbar zu machen, bietet sich reflexives Schreiben an (vgl. Schärli 2013). Denn beim Schreiben hat man ausreichend Zeit seine Gedanken zu sammeln und sie adäquat zu formulieren (vgl. Glesner Fines 2014). Reflektieren benötigt Zeit, die bei mündlichen Äußerungen oft nicht zur Verfügung steht. Moon (2006) bezeichnet deshalb die Aktivität des reflexiven Schreibens auch als „vehicle for reflection“ (ebd., S. 1). In der Literatur herrscht Übereinstimmung darüber, dass Portfolios und Lerntagebücher, sowohl in analoger als auch digitaler Form, sinnvoll für die Initiierung von Reflektierprozessen

sind (vgl. Häcker 2007; Gläser-Zikuda und Hascher 2007). Hilzensauer (2008) betont diesbezüglich, dass (E-)Portfolios und Journals jedoch nur eine „leere Infrastruktur“ (ebd., S. 16) bieten, ohne „eine detaillierte Beschreibung, wie [] Reflexion durchzuführen ist [und] was Reflexion [eigentlich] ist“ (ebd., S. 15). Er schlussfolgert aus seinen Beobachtungen Folgendes: „Für einen institutionellen Kontext ist es nicht zielführend, Portfolios einzuführen ohne dafür konkrete methodisch [sic!] Handlungsanleitungen über die reflexiven Prozesse zu geben“ (ebd., S. 16). Portfolios und andere Möglichkeiten für reflexives Schreiben sind folglich per se keine Reflektierhilfen, sondern schaffen lediglich Raum für die Initiierung von Reflektierprozessen. Damit sie jedoch dazu beitragen, dass sie Reflektierfähigkeiten auch fördern und nicht nur visualisieren, müssen sie didaktisch sinnvoll aufbereitet und beispielsweise durch Feedback oder Instruktionen angereichert werden (vgl. Mansvelder-Longayroux et al. 2007). Portfolios werden in Forschungsarbeiten einerseits als Diagnoseinstrument, zur Erfassung und zum Sichtbarmachen von Lernprozessen verwendet, zum anderen als Bestandteile von Ausbildungsprogrammen oder der Förderung spezieller Kompetenzen (vgl. Landmann und Schmitz 2007). In der hier vorliegenden Studie werden die Blogbeiträge im Sinne einer Portfolioarbeit für ersteres eingesetzt. Sie dienen als reine Initiierung von Reflektierprozessen ohne diese direkt zu fördern. Erst durch Erhalt der Reflektiermethoden und durch das daran angepasste Feedback formt das Bloggen einen Bestandteil des gedachten Förderprogramms der Reflektierfähigkeiten (vgl. Kapitel 3.3.4).

Unterstützung während des Vorgangs des reflexiven Schreibens kann sehr unterschiedlich aussehen. Es gibt die Möglichkeit Reflexionsimpulse zu geben, den Schreibprozess durch Feedback zu begleiten oder konkrete Anleitungen fürs Reflektieren bereitzustellen. Nach Herzog (1995, S. 267) und Hilzensauer (2017, S. 13) sind Unterstützungsangebote essentielle Rahmenbedingungen, um die Förderung von Reflektierprozessen zu ermöglichen. Im Folgenden wird auf die drei Möglichkeiten genauer eingegangen.

Eine geeignete Technik um die Reflektierfähigkeit zu fördern, ist das sogenannte Prompting, worunter das Anregen zum Nachdenken unter Zuhilfenahme von Reflexionsimpulsen verstanden wird (vgl. van den Boom et al. 2007). Mehrere Studien belegen den positiven Einfluss von *reflection prompts* in Hinblick auf das Reflektieren (vgl. Hung 2014; Moon 2004). Damit Lernende Reflektieren nicht mit reinem Berichterstaten verwechseln und sich beim Schreiben nicht nur auf die Form sondern auch den Inhalt konzentrieren, schlägt Glesner Fines (2014, S. 2) folgende Impulse vor:

- Lernende sollen ihre Erfahrung oder Handlung zunächst beschreiben
- Die Handlungen sollen in Hinblick auf allgemein akzeptierte Normen evaluiert werden
- Die eigene Entscheidung soll mit Alternativen verglichen werden

- Motive, Werte oder Annahmen sollen identifiziert werden, die zu den eigenen oder den Handlungen anderer führen
- Die Entscheidungen sollen in Hinblick auf höhere Theorien eingeordnet werden
- Lernende sollen verstehen, wie künftige Entscheidungen durch die eigenen Erfahrungen und Reflexionen beeinflusst werden

Bei ihren vorgestellten Reflektierimpulsen werden überwiegend Reflexivitätsaspekte angesprochen, sodass man insgesamt festhalten kann, dass bei der Verwendung dieser Tipps eher das Reflektieren über das eigene Handeln gefördert werden soll. Glesner Fines (2014) Verständnis vom Reflektieren bezieht sich somit überwiegend auf das Konzept eines *reflective practitioners* nach Schön (1983). In der hier vorliegenden Arbeit wird der Operator *reflektieren* verwendet, sodass das Prompting entsprechend berücksichtigt wird. Das Verständnis von Reflektieren wird vorab mit ihnen geklärt und sie erhalten eine Begriffsdefinition dafür (vgl. Anhang, S. A1).

Eine weitere förderliche Methode ist das Feedbackgeben. So kommt Cendon und Bischoff (2014, S. 36) zu dem Ergebnis, dass Feedback wichtige Impulse für Reflektierprozesse gibt. Durch Rückmeldungen wird beispielsweise der Zugang zu Geschehnissen aus unterschiedlichen Perspektiven ermöglicht, die vorher nicht ersichtlich sind (vgl. Bain et al. 2002, S. 173). Die Strukturierung der eigenen Eindrücke und Erfahrungen wird durch das regelmäßige Feedbackgeben eingeschult (vgl. Hänssig 2010, S. 156). Van den Boom et al. (2004) konnten durch eine Studie aufzeigen, dass die Kombination aus Feedback durch Tutoren und das Einsetzen von *reflection prompts* die Fähigkeit des Reflektierens besonders fördern. Eine weitere Studie von Chen et al. (2009) zeigt, dass Prompts, die vor allem die hohen Reflexions- und Reflexivitätsstufen ansprechen, zusammen mit Beobachtungshinweisen von Experten das Erreichen hoher Reflektierstufen am ausschlaggebendsten beeinflussen. Feedback von Kolleginnen und Kollegen hat hingegen keinen besonderen Einfluss. Auch in dieser Arbeit wird auf das Feedbackgeben nicht verzichtet, da es als wichtiger Bestandteil von Portfolioarbeit angesehen wird. Die Rückmeldungen dabei beziehen sich jedoch stets auf die Anwendung der erhaltenen Reflektierhilfen, da deren Wirksamkeit im Fokus der Untersuchung steht. Sie dienen folglich dazu, nicht selbst das Reflektieren zu fördern, sondern die Schülerinnen und Schüler stets an die erlernten Methoden zu erinnern und ihnen Tipps für deren Umsetzung zur Förderung der Reflektierfähigkeiten zu geben.

Konkrete Anleitungen fürs Reflektieren stellen die dritte Möglichkeit der Förderung von Reflektierfähigkeiten dar. In diesem Zusammenhang bezeichnet Wyss (2008) das Sich-Gedankenmachen über die eigenen Handlungen als „geschlossene Reflexion“ (ebd., S. 6), sobald Materialien dafür zu Hilfe gegeben werden, die „theorie- und praxisorientiert auf die wesentlichen Elemente der Reflexion aufmerksam machen“ (ebd.). Vorteile der geschlossenen Reflexion gegenüber der offenen Reflexion, bei der der zu reflektierende Inhalt nicht vorgegeben wird,

sind die Zielgerichtetheit und ein koordinierter Ablauf. Nachteilig ist jedoch die damit verbundene Beschränkung auf spezielle Aspekte, sodass möglicherweise zu wenig Freiraum für kreatives Denken in andere Richtungen gegeben wird (vgl. ebd.). Doch gerade als Einstieg in das Reflektieren ist die geschlossene Reflexion¹⁷ der offenen Reflexion vorzuziehen, um eine Struktur hineinzubringen. In dieser Studie wird der Fokus auf die Förderung von Reflektierfähigkeiten durch konkrete Anleitungen bzw. Strukturierungshilfen gelegt. Bisher gibt es in der Literatur noch kaum Erkenntnisse darüber, welche Anleitungen sich als adäquate Reflektiermethoden herausstellen. Es werden häufig reflektierfördernde Methoden erwähnt, doch selten Belege für die zugeschriebene Eigenschaft gegeben (vgl. Hilzensauer 2008, S. 1).

2.5.2 Methoden zur Förderung von Reflexionsprozessen

Obwohl das Reflektieren als wichtige Fähigkeit angesehen wird und einige Handbücher zur Förderung dieser existieren, wurden für den Unterricht bisher noch keine konkreten Anleitungen bzw. Methoden für das Einüben des Reflektierens entwickelt. Gibbs Reflexivitätszyklus und die Hütemethode von de Bono sind zwei weitverbreitete Methoden zu deren Potenzialen es jedoch noch kaum empirische Untersuchungen gibt. Ersteres findet häufig Einsatz in der Förderung des reflexiven Handelns, die Hütemethode orientiert sich eher an reflexiven Denken. Im Folgenden werden die beiden Reflektierhilfen vorgestellt, die in dieser Studie als Untersuchungsgegenstand gewählt wurden. Außerdem werden Gründe dafür genannt, warum sie als geeignetes Instrument zur Förderung der Reflexion und Reflexivität eingestuft werden.

2.5.2.1 Gibbs Reflexivitätszyklus

Nach Hilzensauer (2008, S. 1) sind reflexive Lernmethoden ein wichtiger Bestandteil, um lebenslanges Lernen zu ermöglichen. Um Kompetenzen lebenslang weiter auszubilden, müssen sich Lernende explizit mit ihrem eigenen Handeln auseinandersetzen und es ständig analysieren und neu bewerten. Reflexivität nach dem Verständnis von Brendel (2017) ist für diesen Vorgang unerlässlich (vgl. ebd.). Deshalb ist es nicht verwunderlich, dass es zahlreiche Anleitungen gibt, ein *reflective practitioner* nach Schön (1983) zu werden. Das Problem vieler methodischer Hilfestellungen ist jedoch, dass die theoretischen Grundlagen verborgen bleiben und oft unklar ist, welchem Verständnis von Reflexion der Anleitung zu Grunde liegt (vgl. Hilzensauer 2008, S. 1). Ob die Effekte solcher Methoden auch tatsächlich den Gewünschten entsprechen, nämlich der Förderung von Reflexivität, lässt sich folglich nur schwer bis kaum nachvollziehen. Deshalb ist es wichtig, vor der Verwendung einer Reflexivitätshilfe genaues-

¹⁷ *Reflexion* wird nur in diesem Abschnitt als Oberbegriff für das Reflektieren, das aus der Reflexion und der Reflexivität besteht, gesehen, da Wyss (2008) ursprünglich diesen Begriff verwendet.

tens zu prüfen, ob die entsprechende Methode auf einer bestimmten Theorie aufbaut und welches Verständnis von Reflektieren als Grundlage dafür herangezogen wird. Eine Methode, die explizit darauf abzielt, reflexive Praxis zu fördern (vgl. Jasper 2013, S. 77) und deren theoretische Fundierung nachvollzogen werden kann, ist der Reflexivitätszyklus von Graham Gibbs (1988)¹⁸, der in dieser Studie Anwendung findet. Im Folgenden wird näher auf dessen Entwicklung und Einsatz eingegangen.

Der Reflexivitätszyklus von Gibbs (1988) basiert auf David Kolbs *learning cycle* (1984). Mithilfe des Lernzyklus beschreibt Kolb vier Phasen des Lernens, die für ihn Bestandteil davon sind. Wissen und Weisheit werden nach seiner Vorstellung durch die Transformation von Erfahrung innerhalb eines iterativen Prozesses geschaffen. Die vier Phasen bezeichnet Kolb als „learning from concrete experiences (feeling)“, „learning from reflective observation (watching)“, „learning from abstract conceptualization (thinking)“ und „learning from active experimentation (doing)“ (Kolb 1984, S. 38). Wichtig ist es, den Zyklus vollständig zu durchlaufen, da dadurch Lernen aufgebaut und verbessert wird. Grundlage für Kolbs Lernzyklus und auch für Gibbs Reflexivitätszyklus ist das Konzept des erfahrungsbasierten Lernens, das besagt, dass sinnstiftendes Lernen erst durch eine praktische Auseinandersetzung mit einem Lerngegenstand möglich wird (vgl. Stangl 2019). Erfahrungen außerhalb künstlich erzeugter Lernumgebungen sind dabei unerlässlich, da der Lernende selbst als Handelnder im Mittelpunkt steht. Für Kolb (1984, S. 20) stellt sein Konzept des erfahrungsbasierten Lernens eine andere Perspektive auf den Lernprozess dar und unterscheidet sich deutlich von behavioristischen Lerntheorien. Obwohl Kolb wenig über den Bestandteil des Reflektierens in seinem Lernzyklus sagt, ist es dennoch ein wichtiger Aspekt im Lernprozess (vgl. Boud et al. 1993, S. 99). Gibbs (1988) greift die Bedeutung der Reflexivität genauer auf, da Erfahrung nur sinnstiftend sei, wenn sie genauer hinterfragt werde: „Learners must selectively reflect on their experience in a critical way rather than take experience for granted and assume that the experience on its own is sufficient“ (ebd, S. 19). Das tiefgründige Nachdenken, das sowohl mit Gibbs Reflexivitätszyklus als auch mit Kolbs Lernzyklus angeregt werden soll, entspricht Schöns (1983) Begriffsverständnis einer *reflection on action* und folglich der Reflexivität nach dem Verständnis von Brendel (2017), das für diese Studie übernommen wurde.

Gibbs Reflexivitätszyklus, der als Methode zur Förderung der Reflexivitätsfähigkeit im Rahmen dieser Untersuchung ausgewählt und leicht adaptiert wird, ist eine Modifizierung des iterativen Zyklus von Kolb mit der Ergänzung von praktischen Hinweisen für die konkrete Umsetzung und einer stärkeren Einbeziehung von Emotionen. Der Reflexivitätszyklus wurde erstmals in Gibbs Publikation *Learning by doing* (1988) vorgestellt. Er besteht aus sechs einzelnen Phasen, die durchlaufen werden sollen, um ein vorangegangenes Ereignis zu reflektieren (vgl.

¹⁸ Gibbs *reflective cycle* im Original

Tabelle 8). Jede Phase besteht aus mehreren Hilfsfragen, die dabei helfen, strukturiert und richtungsgelenkt über das eigene Verhalten nachzudenken und mögliche Alternativen zu bewerten.

Tabelle 8 Phasen des reflective cycle nach Gibbs (2013, S. 49 f.)

Description: What happened? Don't make judgements yet or try to draw conclusions; simply describe.
Feelings: What were your reactions and feelings? Again don't move on to analysing these yet.
Evaluation: What was good or bad about the experience? Make value judgements.
Analysis: What sense can you make of the situation? Bring in ideas from outside the experience to help you. What was really going on? Were different people's experiences similar or different in important ways?
Conclusions (general): What can be concluded, in a general sense, from these experiences and the analyses you have undertaken? Conclusions (specific): What can be concluded about your own specific, unique, personal; situation or way of working?
Personal action plans: What are you going to do differently in this type of situation next time? What steps are you going to take on the basis of what you have learnt?

Die erste Phase bezeichnet Gibbs (2013, S. 49) als *description*. Wie der Name bereits sagt, soll hierbei die Situation bzw. Handlung, über die reflektiert werden soll, beschrieben werden. Dazu gehören die Vorstellung der beteiligten Personen, die Handlung selbst und der Grund für die eigene Involvierung. Wichtig ist, dass diese Phase kurzgehalten wird, da sie selbst noch keine eigentliche Reflexivität beinhaltet. Dennoch wird sie benötigt, um nachfolgende Überlegungen nachvollziehen zu können. Sie bildet somit den Rahmen für die tatsächliche Reflexivität. Phase Zwei bezieht sich auf die Gedanken und Emotionen, die während der Handlung auftreten. Eine Wertung der Gefühle soll in dieser Phase noch nicht stattfinden, sondern sie sollen lediglich beschrieben werden. Außerdem soll Auskunft darüber gegeben werden, inwieweit bestimmte Gedanken oder Gefühle die Handlung beeinflusst haben. Die dritte Phase, die *evaluation* genannt wird, soll es ermöglichen, die konkrete Situation zu bewerten. Hierzu zählen sowohl die positiven als auch negativen Aspekte der eigenen Handlung und die Bewertung der Verhaltensweisen anderer Beteiligter (vgl. ebd.). Die bisherigen Phasen beziehen sich darauf, was konkret passiert ist. Die nachfolgenden drei Phasen dienen dazu, die Handlung zu analysieren und zu überlegen, wie künftig mit einer ähnlichen Situation umgegangen werden soll. Phase Vier, die Gibbs (2013, S. 49) als *analysis* bezeichnet, ist diejenige Phase, die am ausführlichsten behandelt werden sollte. An dieser Stelle sollen Erkenntnisse aus der Situation

gewonnen werden und die einzelnen Aspekte der vorherigen Phasen daraufhin untersucht werden. Gibbs (2013, S. 50) weist daraufhin, dass in dieser Phase gerne auf wissenschaftliche Erkenntnisse zurückgegriffen werden kann, um Handlungen erklären bzw. aufgetretene Probleme begründen zu können. Als nächster Schritt folgt die Phase *conclusion*, in der die bisherigen Überlegungen zusammengefasst und alternative Handlungsmöglichkeiten in Bezug auf die gesammelten Erkenntnisse vorgestellt werden sollen (vgl. ebd.). Gibbs unterscheidet hierbei eine allgemeine und eine spezifische Schlussfolgerung. Ersteres bezieht sich auf allgemeine Erkenntnisse und letzteres auf sich selbst. Abschließend wird in der Phase *action plan* ein Plan für zukünftige, ähnliche Ereignisse erstellt. Es wird geklärt, welches Verhalten man beibehalten und was man wie ändern würde (vgl. ebd.).

Der Reflexivitätszyklus ist wie Kolbs Lernzyklus iterativ angelegt, sodass er nicht nur einmalig beim Überdenken einer speziellen Handlung verwendet, sondern fortlaufend zur Hilfe genommen werden soll, wenn die zu reflektierende Situation künftig wieder auftritt. Erst durch das ständige Reflektieren über das eigene Verhalten kann eine dauerhafte Änderung in der Handlung erzielt werden (vgl. Gibbs 1988). Das Bewusstmachen von normalerweise unbewussten Handlungsweisen ist ausschlaggebend, um den Mehrwert eines neuen Verhaltens erkennen und folglich einstudieren und verfestigen zu können. Gibbs Reflexivitätszyklus (1988) ist eine Methode, die dieses Bewusstmachen von Unbewusstem unterstützt. Sie kann sowohl beim Reflektieren über konkrete Situationen als auch beim Reflektieren über verfestigte Verhaltensgewohnheiten verwendet werden: „The experience might not be a single occurrence of an event, but might be continuous and varied in terms of its components“ (Moon 2013, S. 27).

Gibbs Reflexivitätszyklus ist eine auch heute noch weit verbreitete Methode zur Ausbildung von *reflective practitioners* nach Schön (1983). Gibbs ursprüngliches Werk *Learning by doing* von 1988 wurde 2013 von der Oxford Brookes University als Onlineausgabe unter einer Creative Commons Lizenz veröffentlicht, um sein Konzept weltweit frei zugänglich zu machen. Im Vorwort der Onlineausgabe heißt es:

Learning by Doing' has been a hugely influential text in education. In particular the reflective cycle described here has been widely adopted by those studying, practising and teaching the skills of critical reflection. It has become a seminal text for health care professionals, staff developers and higher education teachers, and remains a recommended text on reading lists to this day. (Gibbs 2013, S. 6)

Besonders in der Ausbildung von medizinischen Fachkräften im englischsprachigen Raum wird das Werk und der Reflexivitätszyklus häufig verwendet (vgl. Quinn 2000, Rana und Upton 2013). Trotz des vielfachen Einsatzes der Methode zur speziellen Förderung der Reflexivität gibt es nur wenig Forschungsarbeiten über das tatsächliche Potenzial der Methode (vgl. 2.7). Diese Studie verfolgt deshalb das Ziel, die in der Literatur vielfachen prognostizierten Erfolge

hinsichtlich des reflexiven Handelns durch die Anwendung von Gibbs Reflexivitätszyklus praxisorientiert zu testen. Im deutschsprachigen Raum findet die Methode bisher kaum Anwendung (vgl. Schmal 2018), im Kontext einer BNE ist der Autorin kein Einsatz bekannt.

Gibbs Reflexivitätszyklus eignet sich für die Förderung der Reflexivitätsfähigkeit, da es praxisorientiert ist und sich nicht nur auf das Reflektieren bei Lernprozessen bezieht, sondern allgemein gehalten ist (vgl. Hilzensauer 2008, S. 6). Der Fokus der Methode liegt eindeutig auf dem Reflektieren über die Handlung, sodass das Konzept des reflexiven Handelns nach Schön (1983) als theoretischer Rahmen dient. Die Inhalte des Zyklus basieren auf der erfahrungsbauierten Lerntheorie und Grundlage für Gibbs Reflexivitätszyklus bildet Kolbs Lernzyklus. Aufgrund der eben genannten Eigenschaften wird die Methode als geeignete Methode zur Förderung von Reflektierfähigkeit in dieser Studie eingeordnet. Ob sie tatsächlich den Teilbereich Reflexivität des Reflektierens besonders fördert, soll dabei herausgefunden werden. Da Gibbs Reflexivitätszyklus im Original nur in englischer Sprache existiert, wird er für die Studie übersetzt und an den Kontext der BNE angepasst (vgl. Kapitel 3.3.4).

2.5.2.2 Hütemethode nach de Bono

Wie bereits in Kapitel 2.1.1.1 erläutert, ist die Reflexion nach dem Verständnis von Dewey (1933) ein Kernelement des kritischen Denkens. Kraak (o. J.) betont die Wichtigkeit, Schülerinnen und Schülern explizit die Fähigkeit des kritischen Denkens zu lehren. Diese gehört wie Kivunja es ausdrückt, zu den „super skills“ (Kivunja 2015a, S. 4), die essentiell für den Erfolg in der Ausbildung, im Beruf und im alltäglichen Leben des 21. Jahrhunderts sind. Edward de Bono, Begründer des lateralen und parallelen Denkens, weist darauf hin, dass Denkfertigkeiten am effektivsten gefördert werden, wenn sie explizit gelehrt werden und nicht indirekt Verwendung finden (vgl. Bono 1992b, S. 7). Er stützt seine Erkenntnis auf seine eigene langjährige Erfahrung in diesem Bereich. Neben dem konkreten Vorstellen von Methoden zur Förderung kritischen Denkens, ist er der Meinung, dass einfache Methoden weitaus effektiver sind, als komplizierte, schwer zu verstehende Methoden, die Menschen eher verwirren als Unterstützung im Denkprozess zu geben (vgl. Bono 1992b, S. 6). Er entwickelte deshalb eine simple Methode zur Förderung des kritischen und kreativen Denkens (vgl. Marrapodi 2003, S. 23), die sich an der Metapher des Aufsetzens von Hüten orientiert. Seine Sechs-Hüte-Methode, die er erstmals im Jahr 1985 ausführlich in seinem Werk *Six Thinking Hats* vorstellte, ist inzwischen international bekannt und wird sowohl zur Schulung von Führungskräften als auch im Bildungsbereich genutzt (vgl. Bono 2005, S. 11). Im Folgenden soll kurz die Methode selbst vorgestellt werden, bevor auf deren Grundlagen eingegangen und abschließend begründet wird, warum diese Methode zur Förderung der Reflexion im Rahmen dieser Dissertationsarbeit gewählt wurde.

Die *Six Hats Method* ist eine Methode, bei der mehrere Personen zur selben Zeit ein Thema von bestimmten Perspektiven aus untersuchen (vgl. Bono 1992b). Wie der Name bereits sagt, besteht die Methode aus sechs unterschiedlichen Denkperspektiven, denen symbolisch verschieden farbige Hüte zugeordnet sind. Der weiße Hut steht für die objektive Sichtweise, der rote Hut für die emotionale Sichtweise, der schwarze Hut für die pessimistische Sichtweise, der gelbe Hut für die optimistische Sichtweise und der grüne Hut für die kreative Sichtweise. Der blaue Hut hat eine Kontrollfunktion für den Denkprozess und steht über den anderen Hüten (vgl. ebd.). Der Ablauf der Methode ist nicht fest vorgegeben. Wichtig ist, dass alle Teilnehmenden gleichzeitig jeweils den selben Hut tragen und sie ein Thema bzw. Problem somit immer aus derselben Perspektive betrachten. Nacheinander sollen die unterschiedlichen Hüte getragen und somit eine multiperspektivische Sichtweise auf ein Thema ermöglicht werden. Der blaue Kontrollhut kann von einem Gruppenleiter übernommen werden, der die Ergebnisse der einzelnen Hütaperspektiven festhält und kontrolliert, dass beim Hütetragen auch nur die vorgegebenen Sichtweisen eingenommen werden und die Methode strukturiert abläuft (vgl. ebd.). Im Gegensatz zur Methode der Debatte oder strukturierten Kontroverse werden bei der Hüte-Methode keine Argumente gegenübergestellt und am Ende nach einer gemeinsamen Lösung gesucht, sondern unterschiedliche Ansichten auf ein Thema aufgedeckt und bewusstgemacht. Sie beabsichtigt nicht, am Ende Sichtweisen als falsch oder richtig zu interpretieren, sondern den Teilnehmenden zu verdeutlichen, dass es je nach unterschiedlichen Voraussetzungen unterschiedliche Meinungen zu Themen geben kann und man all diese bei den eigenen Überlegungen beachten muss. De Bono bezeichnet die Denkart, bei der eine Gruppe jeweils gleichzeitig eine Denkhaltung einnimmt, als paralleles Denken (i.Org. *parallel thinking*; vgl. Bono 2000, S. 4). Die klare Trennung der einzelnen Denkrichtungen ist dabei ein wichtiges Element und muss durch den Träger des blauen Hutes stets eingefordert werden.

Ausgangspunkt für die Entwicklung der Hüte-Methode war für de Bono (2000, S. 1 ff.) ein nötiges Umdenken im Denkprozess, da unser Denksystem nicht mehr zeitgemäß sei. Er ist der Meinung, dass unser Denksystem, das auf Beurteilungs- und Gegensatzdenken basiert und durch Platon, Sokrates und Aristoteles geprägt wurde, zwar die Gesellschaft weitergebracht und Entwicklungen vorangetrieben hat, es jedoch in der heutigen Zeit ungeeignet sei. Unser traditionelles Denken sei nicht falsch, doch unbefriedigend, da es nicht adäquat auf Mehrdeutigkeiten, rasante Veränderungen und sich ergebende Potenziale reagieren kann (vgl. ebd.). Seiner Meinung nach sind intelligente Menschen nicht automatisch gute Denker. Zwar besitzen sie eine ausgeprägte Fähigkeit, ihre Meinungen begründet darzustellen und andere Ansichten mit Gegenargumenten zu widerlegen, da sie dadurch ihre Intelligenz zur Geltung bringen können. Doch dadurch haben sie oftmals Schwierigkeiten, neue Ideen zu akzeptieren oder nach Alternativen zu suchen (vgl. Bono 2005, 16 f.). Sie müssen ihre Auffassung, immer Recht zu haben, überdenken und dies erfordert es, sich dieser sogenannten „Intelligenzfalle“

(ebd.) bewusst zu werden und gegen sie zu arbeiten. Folgendes Zitat veranschaulicht de Bono's Intention der Sechs-Hüte-Methode:

Western thinking is concerned with 'what is', which is determined by analysis, judgement and argument. That is a fine and useful system. But there is another whole aspect of thinking that is concerned with 'what can be', which involves constructive thinking, creative thinking and 'designing a way forward [...] Yet the basic tradition of Western thinking (or any other thinking) has not provided a simple model of constructive thinking. That is precisely what the Six Hats method (parallel thinking) is all about. (Bono 2000, 2 f.)

Neben dem Ausgangspunkt, das Denksystem aktualisieren zu wollen, liegt der Hüte-Methode eine weitere philosophische Annahme zugrunde. De Bono (1992c, S. 8) ist der Ansicht, dass dem Menschen beim Versuch praktisch zu denken, drei Schwierigkeiten begegnen, die er als Emotionen, Hilflosigkeit und Verwirrung identifiziert. Die emotionale Schwierigkeit besteht darin, dass sich Menschen oft beim Denken durch ihr Bauchgefühl oder durch Vorurteile beeinflussen lassen und dadurch nicht in der Lage sind, allumfassend nachzudenken. Die Hilflosigkeit drückt sich in Form des Gefühls der Unzulänglichkeit aus. Diese entsteht, wenn sich manche Menschen mit dem Problem auseinandersetzen, dass sie nicht wissen, wie sie über ein bestimmtes Thema denken sollen. Die dritte Schwierigkeit der Verwirrung entsteht dadurch, dass man beim Nachdenken versucht, an alles gleichzeitig zu denken und dadurch kein zufriedenstellendes Ergebnis erzielen kann, sondern ein Durcheinander in der eigenen Gedankenwelt schafft (vgl. ebd.).

Mithilfe der Hüte-Methode sollen diese Probleme umgangen werden (vgl. Bono 1992c, S. 8). Dadurch, dass den Teilnehmenden der Hüte-Methode stets vorgegeben wird, welchen Hut sie gerade tragen und aus welcher Perspektive sie somit das Thema betrachten sollen, kann verhindert werden, dass Emotionen den Gedankenprozess stören. Gefühle werden im Rahmen der Hüte-Methode berücksichtigt, jedoch setzen sie nicht unkontrolliert ein, sondern nur an der gewünschten Stelle, d.h. beim Tragen des roten Hutes. Das Problem der Hilflosigkeit bei der Konfrontation mit schwierigen, undurchsichtigen Themen kann durch eine geplante Vorgehensweise gelöst werden. Die Hüte-Methode bietet dafür einen geeigneten Rahmen und ermöglicht es, Schritt für Schritt im eigenen Denkprozess strukturiert voranzukommen (vgl. ebd.). Um der Verwirrung entgegenzuwirken hilft das parallele Denken. Verwirrung im Denkprozess entsteht, wenn wir an alles gleichzeitig denken wollen und uns dann in unseren Gedanken verlieren. Beim Anwenden der Hüte-Methode konzentriert man sich gemeinsam mit anderen bewusst nacheinander auf unterschiedliche Sichtweisen auf ein Thema und kann somit einem roten Faden während des Denkprozesses folgen, durch den man geordnet verschiedene Denkrichtungen berücksichtigt (vgl. ebd.).

Geht man von der Gründungszeit der Hüte-Methode aus, entsteht zunächst der Eindruck, dass sie möglicherweise veraltet sei und bereits modifiziert wurde. Dem ist aber nach aktuellem Wissensstand der Forscherin nicht so. Im Gegenteil, die Methode von de Bono wird auch heute

noch weltweit für unterschiedliche Zwecke eingesetzt, sowohl im Bildungsbereich als auch in der freien Wirtschaft. De Bono berichtet in seiner aktuellsten Auflage seines Werks *Six Thinking Hats* von 2017 davon, dass seine Methode in populären Unternehmen wie Shell, DuPont oder der NASA verwendet wird, während gleichzeitig Schulen und Universitäten im Rahmen ihres Bildungsauftrages davon profitieren (vgl. Bono 2017, Preface). Auch im deutschen Raum ist die Methode weit verbreitet und wird häufig als geeignete Unterrichtsmethode angewendet, um Schülerinnen und Schülern unterschiedliche Sichtweisen auf ein Thema aufzuzeigen und einem linearen Denken entgegenzuwirken¹⁹.

Die Tatsache, dass die Hüte-Methode bereits seit mehreren Jahrzehnten erfolgreich in unterschiedlichen Bereichen Anwendung findet und bisher nicht überarbeitet wurde, rechtfertigt die Auswahl dieser Methode zur Förderung der Reflexion von Jugendlichen. Reflexion ist eng mit dem kritischen Denken verwandt (vgl. Kapitel 2.1.2) und kann durch den Aspekt der Suche nach alternativen Lösungen sogar als Erweiterung davon gesehen werden. De Bono betont die Wichtigkeit, neben dem kritischen Denken, das aus der Analyse, Urteilsbildung und Diskussion besteht, auch die kreative und konstruktive Seite des Denkens nicht zu vernachlässigen (vgl. Bono 2005, 20 f.). De Bono gilt als einer der bedeutendsten Lehrer²⁰ des parallelen Denkens und sein ursprüngliches Werk *Six Thinking Hats* wurde inzwischen in mehr als 20 Sprachen übersetzt (vgl. Kivunja 2015b, S. 381). Ein weiterer Grund, der den Einsatz der Methode für diese Forschung stützt, ist, dass die Methode sowohl für Kinder im Grundschulalter als auch für Führungskräfte geeignet ist (vgl. Bono 1992a, S. 161) und somit altersunabhängig eingesetzt werden kann.

Abschließend seien noch auf die Vorteile dieser Methode zur Förderung der Reflexion im Bereich der BNE gegenüber anderen kooperativen Methoden zu nennen. Wie bereits erwähnt, basiert die Hüte-Methode auf der Form des parallelen Denkens (vgl. Bono 2000, S. 4). Der große Vorteil gegenüber weiteren, etablierten Unterrichtsmethoden, wie beispielsweise der Pro-Kontra Debatte oder der Fishbowl-Methode, ist der, dass bei der Hüte-Methode jede Perspektive auf ein Thema gemeinsam untersucht und besprochen wird und somit nicht verlangt wird, jemanden von einer bestimmten Ansicht zu überzeugen. Vielmehr ist es das Ziel, die Komplexität eines Themas durch unterschiedliche Perspektiven darauf nachvollziehen und Gründe für bestimmte Ansichten verstehen und bewerten zu können. Wie später noch beschrieben wird, ist die Multiperspektivität auf ein Thema ein bedeutender Aspekt, um auf hohem Niveau über ein Thema reflektieren zu können. Voraussetzung dafür ist natürlich, dass die Methode richtig eingesetzt wird und die Hüte nicht an einzelne Schülerinnen und Schüler

¹⁹ Zum Beispiel listet die Bundeszentrale für Politische Bildung auf ihrer Homepage die Hütemethode in ihrer Methodendatenbank für den Unterricht auf und auch in den Diercke Methoden findet sich die Methode wieder.

²⁰ Er hielt bereits in mehr als 50 Ländern Vorträge und lehrte u.a. an den Universitäten in Oxford, P5, Cambridge und Harvard. Zudem werden seine Denkstrategien international in Form von Ausbildungskursen angeboten (Bono 2005, S. 11).

verteilt werden, die diese den anderen präsentieren sollen. Dies entspricht nicht der Bonos Grundidee und dürfte korrekterweise auch nicht als Sechs-Hüte-Methode bezeichnet werden. Wird der Fokus auf das parallele Denken gelegt, können mit dieser Methode die unterschiedlichen Ansichten auf nachhaltige Themen beteiligter Personen begründet werden. Das Erkennen der Komplexität und das Einnehmen und Akzeptieren unterschiedlicher Perspektiven ist im Bereich der BNE laut der Gestaltungskompetenz nach de Haan (2008) und den Ergebnissen mehrerer Autoren (vgl. Grobbauer und Thaler 2010, 125 ff.; Ohl und Klebel 2012, 75 ff.) unabdingbar.

Ein konkreter Einsatz im Unterricht könnte wie folgt aussehen: In einem ersten Schritt wird gemeinsam die Hüte-Methode durchgespielt, in einem zweiten Schritt und in Anlehnung an den blauen Hut werden anhand des Konzepts des Nachhaltigkeitsdreiecks mehrere Perspektiven auf das Thema überlegt, die sich durch das Bewusstsein bzw. der Gewichtung unterschiedlicher Hüte ergeben. So kann auch die eigene Meinung auf ein Thema untersucht und festgestellt werden, warum die entsprechende persönliche Wertung existiert. Die Abhandlung der unterschiedlichen Hüte macht die Schülerinnen und Schüler darauf aufmerksam, welche Hüte sie unbewusst bevorzugt aufsetzen und auf welche sie künftig bei ihrer persönlichen Urteilsbildung mehr achten müssen. Die Hüte können als Erklärung für unterschiedliche Ansichten herangezogen werden. Ein Umweltaktivist trägt beispielsweise beim Thema Umweltschutz hauptsächlich den positiven Hut, während ein Unternehmer im Bereich der traditionellen Stromproduktion sich mehr auf negative Aspekte konzentriert und die Vorteile des Umweltschutzes möglicherweise bewusst verdrängt oder sie für sich persönlich weniger stark gewichtet.

Die Bonos Sechs-Hüte-Methode kann angepasst am Konzept des Nachhaltigkeitsdreiecks zum Einsatz kommen und dient als Grundlage dafür, die eigene Perspektive auf ein Thema erklären und bewerten zu können. Außerdem hilft sie dabei, weitere Perspektiven auf Themen aufzudecken und die Gründe dafür nachvollziehen zu können. Wie genau die Methode in dieser Studie zum Einsatz kommt, wird in Kapitel 3.3.4 vorgestellt.

2.6 Das digitale Medium Blog

2.6.1 Definition und Klassifikation von Medien

Medien gehören zum Unterricht und spielen bei der Unterrichtsplanung und Auswahl von geeigneten Methoden eine wichtige Rolle. Nach Meyer (1987) kann „[die] begründete Auswahl und der geschickte Einsatz von Medien [...] den Unterricht entscheidend befruchten“ (ebd., S. 151). Für ihn stellen Medien „‘tiefgefrorene‘ Ziel-, Inhalts- und Methodenentscheidungen“ (ebd.) dar, die durch das methodische Agieren von Lehrenden und Lernenden „wieder ‘aufge-

taut‘ werden“ (ebd.). Weiterhin ist er der Meinung, dass der Medieneinsatz die Selbstständigkeit von Schülern fördert und der Unterrichtsinhalt dadurch auch mit einer reduzierten Präsenz der Lehrkraft angeeignet werden kann (vgl. ebd.).

Der Medienbegriff wird in der Literatur oft unterschiedlich verwendet. Tulodziecki et al. (2004) unterscheiden zwischen einem umfassenden Begriffsverständnis und einem engeren Medienbegriff:

Im allgemeinen Sinne des Wortes kann man die Form, in der ein Inhalt präsentiert wird, als Medium bezeichnen [...] Der Begriffsgebrauch reicht [allerdings] von dem oben angeführten umfassenden Begriffsverständnis als Erfahrungsform bis zu einer engeren Begriffsauffassung, bei der nur dann von Medien gesprochen wird, wenn potenzielle Zeichen mit Hilfe technischer Geräte gespeichert, übertragen oder verarbeitet und in abbildhafter oder symbolischer Darstellung wiedergegeben werden. Beispiele für Medien in diesem engeren Sinne sind Buch, Arbeits- und Diaprojektion, Film, Video und Fernsehen, [...], Computer und Multimedia. (Tulodziecki et al. 2004, S. 135)

Im Bereich der Bildung werden Medien oftmals als „Objekte, technische Geräte oder Konfigurationen [definiert], mit denen sich Botschaften speichern und kommunizieren lassen“ (Weidemann 2002, S. 46). In der Geographiedidaktik werden Medien als „Träger von subjektiv ausgewählten und gespeicherten Informationen [verstanden]. Im unterrichtlichen Lernprozess haben sie eine Mittlerfunktion zwischen der Wirklichkeit und dem Adressaten/Lernenden“ (Rinschede 1999, S. 101).

Die meisten Definitionen von Medien haben den Aspekt der Kommunikation als gemeinsame Eigenschaft. Nach der Klassifikation nach Pross (1970) kann man Medien in drei Kategorien einteilen. Die primären Medien beinhalten alle körpergebundenen Formen der Kommunikation, wozu die Sprache, Mimik und Gestik und der Tanz gehören. Zu den sekundären Medien gehören Übermittlungstechniken, die nur auf Seiten des Kommunikators ein Gerät benötigen, wie beispielsweise Texte, Bilder oder Musik. Die tertiären Medien schließlich „nennen wir Transporttechniken, die zur Produktion *und* zum Konsum Geräte erfordern (Telegraph, Telefon, Fernschreiber, Radio, Television, Film)“ (Pross 1970, S. 129). Tertiärmedien bezeichnen folglich technische Medien, wozu zum Beispiel Telefone oder Fernseher zählen. Da sich seit der Kategorienbildung von Pross weitere Kommunikationsmedien entwickelt haben, wurde die Klassifikation u.a. durch Faßler (1997) modifiziert: Neben den drei Kategorien existiert für ihn eine vierte Kategorie, die er als „quartäre Medien“ (ebd., S. 117) oder auch „digitale Medien“ (ebd.) definiert. Darunter versteht er die „computerbasierten und -verstärkten Medienbereiche netztechnischer und elektronisch-räumlicher Konsumtion, Information und Kommunikation“ (ebd.). Gemeinsame Eigenschaften der quartären Medien sind die Integration der Digitalisierungstechnik, die Notwendigkeit eines Computers oder Ähnlichem und die Verortung in Netzwerken. Wie bei den tertiären Medien werden auch hier sowohl auf der Produzenten- als auch auf der Konsumentenseite technische Mittel benötigt. Schellmann et al. (2017) untergliedern

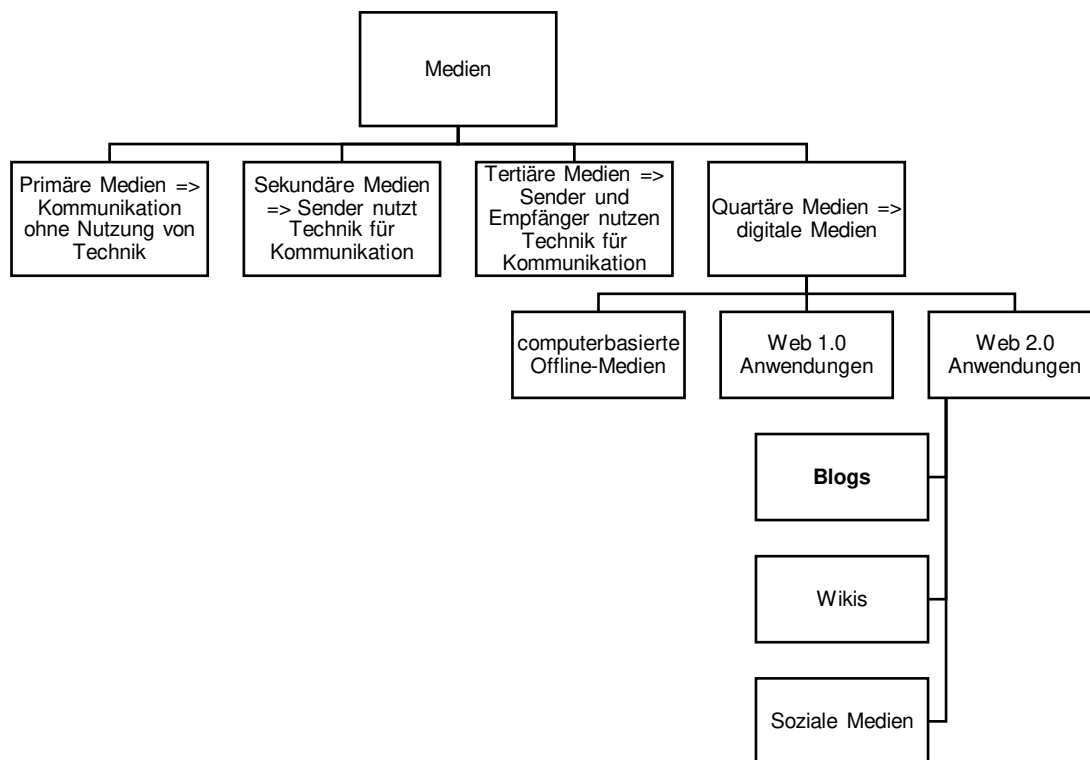
die tertiären Medien erneut in frühe, neuere und neuste tertiäre Medien, letzteres entspricht hierbei Faßlers quartären Medien und beinhaltet das Internet und Smartphones.

Im Bereich des Lernens und Lehrens bezieht man sich seit den letzten zehn Jahren mit dem Medienbegriff mehr und mehr auf Medien, die in Zusammenhang mit der Computertechnologie stehen. Grund ist die zunehmende Verbreitung und Alltagsintegration computerbasierter Techniken und des Internets und die sich stetig weiterentwickelnde Technologie, die neue Kommunikationsmöglichkeiten schafft (vgl. Hüther 2005, S. 346). Autoren verwenden für diesen Typus von Medien häufig die Begriffe *neue Medien* (vgl. Tulodziecki et al. 2004; Zander und Brünken 2006), *computerbasierte Medien* (vgl. Eickelmann und Schulz-Zander 2010; Wild und Möller 2009) oder *digitale Medien* (vgl. Bergmann 2009; Girwidz 2002).

2.6.2 Blogs als Unterkategorie digitaler Medien

Digitale Medien lassen sich – wie Medien im Allgemeinen auch – nach unterschiedlichen Aspekten klassifizieren. Im Folgenden soll eine Klassifizierung nach Nutzungsart vorgenommen werden, um die Positionierung von Blogs innerhalb der digitalen Medien veranschaulichen zu können. Für die Klassifizierung spielt das Web 2.0 eine wichtige Rolle. Der Begriff *Web 2.0* wurde maßgeblich durch Tim O'Reillys Artikel *What is Web 2.0* aus dem Jahr 2005 geprägt. O'Reilly wird seither die Urheberschaft des Ausdrucks zugeschrieben. Brendel und Schrüfer (2013, S. 289) fassen die Hauptmerkmale des Web 2.0 folgendermaßen zusammen: Das Web 2.0 unterscheidet sich vom Web 1.0 dadurch, dass es darin keine klare Trennung zwischen Autorenschaft und Rezipienten gibt, da beide Seiten Inhalte erstellen oder bearbeiten und auf im Web Veröffentlichtes aktiv reagieren können (z.B. in Form von Kommentaren oder durch Bewertungen). Anwendungen innerhalb des Web 2.0 sind Blogs, Plattformen sozialer Netzwerke und Wikis (vgl. ebd.). Einige Anwendungen davon werden oft auch mit dem Begriff soziale Medien umschrieben. „Soziale Medien sind mittlerweile fester Bestandteil der Online-Welt und unter dem Sammelbegriff [...] werden eine Vielzahl ganz unterschiedlicher Web-Angebote verstanden, zu denen vor allem die sozialen Netzwerke wie Facebook, Wikipedia, YouTube und Twitter gehören“ (König et al. 2014, S. 7). Blogs können somit als Unterkategorie von digitalen Medien verstanden werden (vgl. Abbildung 1), da sie zu den Anwendungen des Web 2.0 zählen (vgl. Brendel und Schrüfer 2013, S. 289).

Abbildung 1 Positionierung von Blogs innerhalb der Medienklassifikation (eigener Entwurf; in Anlehnung an Pross (1970), Faßler (1997), Brendel und Schröfer (2013))



Im Gegensatz zu passiven Web 1.0 Medien, wie beispielsweise einer klassischen, statischen Website oder einem im Internet veröffentlichten Foto auf einer Homepage, die dem traditionellen Verständnis eines Mediums als Kommunikationsmöglichkeit bzw. Informationsübermittlung in eine Richtung entsprechen, ermöglichen Web 2.0 Medien dem Empfänger aktive Reaktionen auf das Kommunizierte. Sie können im Rahmen des Web 2.0 sowohl Empfänger als auch Sender gleichzeitig sein und erweitern die Eigenschaften tertiärer Medien nach dem Verständnis von Pross (1970) somit um die Komponente der Internetfähigkeit und den Aspekt der zeitlichen Unabhängigkeit, da auf Kommuniziertes jederzeit und von jedem Ort aus reagiert werden kann. Levine und Vogel (2000) bezeichnen das Web 2.0 als Mitmachmedium, da es im Gegensatz zum vorherigen Web 1.0 nicht mehr nur eine reine statische Informationsquelle ist, sondern inzwischen Interaktionen ermöglicht und zum Mitmachen und Aktivwerden anregt. Genau dies sollte als Potenzial für den Unterricht erkannt und genutzt werden.

2.6.3 Eigenschaften von Blogs

Mit der Verbreitung des Web 2.0 ist auch die Bedeutung und Anzahl von Blogs gestiegen. Zerfaß und Boelter (2005) haben sich in ihrem Werk *Die neuen Meinungsmacher* ausführlich mit der Blogthematik auseinandergesetzt und Eigenschaften von Blogs herausgearbeitet. Ursprünglich verstand man unter einem Blog ein Onlinetagebuch einer Person, das nur mit einer bestimmten Personengruppe geteilt wurde. Von dieser herkömmlichen Bedeutung wurde auch

der Name geprägt. Weblog stellt das Akronym der beiden Begriffe *Web* und *Logbuch* dar, inzwischen wird häufig auch die verkürzte Bezeichnung *Blog* verwendet (vgl. Zerfaß und Boelter 2005, 20 f.). Nach der Definition von Zerfaß und Boelter (2005) sind Blogs „einfache, in ihrem Funktionsumfang meist reduzierte Web-Content-Management-Systeme (Web-CMS), d.h. webbasierte Anwendungen zum Erstellen, Verwalten und Veröffentlichen von Inhalten im World Wide Web“ (ebd., S. 36). Die Definition sagt nichts darüber aus, wie viele Personen daran beteiligt sind, welche thematischen Inhalte vorgestellt werden und in welchem Format dies geschieht. Blogs können sich deshalb hinsichtlich der Themen, der Art des zur Veröffentlichung gewählten Mediums und der potenziellen Zielgruppe unterscheiden (vgl. Koller und Alpar 2008, S. 19). In der Regel wird in Blogs meist ein Hauptthema behandelt, welches in Unterkategorien eingeordnet werden kann. Der Vorteil eines Blogs gegenüber einem nicht-digitalen Heft ist der, dass mehrere Medien, wie Text, Fotos, Tonaufnahmen oder Videos, integriert werden können und die Möglichkeit der Interaktion in Form von Kommentaren oder Verlinkungen besteht (vgl. Matzen 2014, S. 93 ff.).

Für den Unterricht bietet der Einsatz von Blogs viele Chancen. Zum einen können Blogs verwendet werden, um die Organisation zu vereinfachen, zum anderen können Blogs als Kommunikationskanal oder Werkzeug zur Förderung von Kompetenzen fungieren (vgl. Sauter und Sauter 2013, S. 114 f.). Weiterhin kann ein Blog auch als Erhebungsinstrument für Lehrende und als Instrument zur Initiierung von Reflektierprozessen für Lernende genutzt werden (vgl. Stocker 2007, S. 101). Für die Bestimmung von Reflektierfähigkeiten werden häufig schriftlich festgehaltene Ergebnisse aus Portfolios (vgl. Eysel 2006; Leonhard 2008), Berichten (vgl. Hatton und Smith 1995; Korthagen 2002) oder Lerntagebüchern (vgl. Heinzel und Brencher 2008) verwendet. Auch Blogs haben eine ähnliche Funktion der Dokumentensammlung und geben die Möglichkeit Gedanken schriftlich festzuhalten. E-Portfolios, die digitale Komponente von klassischen Portfolios sind nach Tzeng et al. (2015) „a personal digital collection of information describing and evidencing a person's learning, career, experience and achievements“ (ebd., S. 1). Diese Eigenschaft kann auch Blogs zugeschrieben werden, weshalb sie ein ideales Instrument zur Erstellung von E-Portfolios darstellen (vgl. Richardson 2011, S. 45). In der Literatur herrscht Übereinstimmung darüber, dass (E-)Portfolios sinnvoll für die Initiierung von Reflektierprozessen sind (vgl. Gläser-Zikuda 2007; Häcker 2007). Damit Portfolios jedoch dazu beitragen, dass sie Reflektierfähigkeiten auch fördern und nicht nur visualisieren, müssen sie didaktisch sinnvoll aufbereitet und beispielsweise durch Feedback oder Instruktionen angereichert werden (vgl. Mansvelder-Longayroux et al. 2007). Portfolios werden in Forschungsarbeiten einerseits als Diagnoseinstrument, zur Erfassung von Lernprozessen verwendet, zum anderen als Bestandteile von Ausbildungsprogrammen oder der Förderung spezieller Kompetenzen (vgl. Landmann und Schmitz 2007). In der hier vorliegenden Studie werden die Blogbeiträge im Sinne einer Portfolioarbeit für ersteres eingesetzt. Sie dienen als reine Initiierung

von Reflektierprozessen, ohne diese direkt zu fördern. Erst durch Erlernen der Reflektierhilfen und durch das daran angepasste Feedback formen die Beiträge einen Bestandteil des gedachten Förderprogramms der Reflektierfähigkeiten (vgl. Mansvelder-Longayroux et al. 2007; Tzeng et al. 2015).

2.7 Forschungsstand

Obwohl das Reflektieren als wichtige Fähigkeit gesehen wird, die in allen Unterrichtsfächern und auch für lebenslanges Lernen von Bedeutung ist, wird der Vermittlung und Förderung dieser Fähigkeit im schulischen Kontext nur wenig Aufmerksamkeit gewidmet (vgl. Killus 2009). Um dem entgegenzuwirken, wird in der hier vorliegenden Studie der Fokus auf die explizite Förderung von Reflektierfähigkeiten gelegt, mit dem Ziel, Lehrkräften Empfehlungen zum gezielten Üben von Reflektieren mit Lernenden geben zu können. Bisher gibt es nur wenige Forschungsarbeiten, die die Förderung von Reflektierfähigkeiten von Schülerinnen und Schülern untersuchen, weshalb das Forschungsfeld innerhalb dieses Bereichs sehr überschaubar bleibt. Überwiegend wird in Hinblick der Thematik des Reflektierens im schulischen Kontext über die Reflektierfähigkeiten von Lehrenden oder künftigen Lehrkräften geforscht, wie beispielsweise in den Arbeiten von Abels (2011), Wyss (2013) oder Kittel und Rollett (2017). Das Reflektieren wird hierbei als wichtiger Teil zur Entwicklung professioneller Handlungskompetenz gesehen. Die wenigen Studien, die Reflektierfähigkeiten von Schülerinnen und Schülern behandeln, sehen das Reflektieren als Bestandteil des selbstregulierten Lernens im Sinne vom Hinterfragen eigener Lernstrategien (vgl. Leidinger 2014). Hier steht nicht das allgemeine Reflektieren im Fokus, sondern die Optimierung eigener Lernprozesse. Das Reflektieren über eigene Meinungen oder eigenes Verhalten, das besonders im Bereich des globalen Lernens von Bedeutung ist und außerhalb von Lernstrategien Anwendung findet, wird selten thematisiert. Im Folgenden möchte ich die Studie von Brendel (2017) detaillierter vorstellen, da sie die Grundlage für diese Arbeit bildet und anschließend auf weitere Studien im Bereich von Reflektierfähigkeiten innerhalb des Geographieunterrichts und innerhalb des Bereichs Globalen Lernens eingehen. In einem nächsten Schritt möchte ich empirische Ergebnisse zur Wirkung von Gibbs Reflexivitätszyklus und der Hütemethode von de Bono auf das Reflektieren oder eine übergeordnete Fähigkeit, bei der Reflektieren eine Rolle spielt, präsentieren. Abschließend werden die beiden Studien vorgestellt, aus denen die beiden ursprünglichen Reflektiermodelle hervorgehen, die im Rahmen dieser Arbeit für die Datenauswertung von Bedeutung sind.

2.7.1 Brendels Dissertationsstudie (2017)

Die wichtigste Forschungsarbeit für die Entwicklung dieser Forschung ist die Dissertationsstudie von Brendel mit dem Titel *Reflexives Denken im Geographieunterricht. Eine empirische Studie zur Bestimmung von Schülerreflexion mithilfe von Weblogs im Kontext Globalen Lernens* (2017). Die Erkenntnisse daraus legen den Grundstein für die nachfolgende Untersuchung. Hauptaugenmerk der Arbeit liegt auf der Bestimmung der Ausprägungsgrade der Reflexionsfähigkeiten von Schülerinnen und Schülern aus vier Oberstufenklassen im Bereich des Globalen Lernens. Neben der Entwicklung eines Stufenmodells zum reflexiven Denken nach dem Verständnis von Dewey (1933) speziell für den Bereich des Globalen Lernens und einer BNE, leitet die Forscherin Strategien zur Förderung von Reflexionsfähigkeiten ab. Brendel setzt sich in ihrer Studie ausführlich mit den unterschiedlichen Begriffsverständnissen von Reflektieren auseinander. Als Grundlage für ihre Untersuchung nimmt sie eine Unterscheidung zwischen Reflexion und Reflexivität vor:

Reflexion bezeichnet ein Nachdenken über Unterrichtsinhalte, Überzeugungen oder Vorstellungen und geht zurück auf das Konzept des reflexiven Denkens nach Dewey (1997).

Reflexivität bezeichnet ein Nachdenken über eigene Handlungen oder Handlungsalternativen und fußt auf dem Konzept der reflexiven Praxis nach Schön (1983). (Brendel 2017, S. 19; H.i.O.)

Zum Erfassen der Daten werden in der Studie Blogbeiträge verwendet. Die Schülerinnen und Schüler verfassen Texte, in denen sie über die Inhalte des Unterrichts reflektieren sollen. Die Auswertung der einzelnen Beiträge erfolgt mittels der qualitativen Inhaltsanalyse nach Mayring. Die Kategorien für die Auswertung der Blogbeiträge werden induktiv aus dem Stufenmodell von Bain et al. (1999) abgeleitet und deduktiv durch das Material modifiziert. Ergebnis der Auswertung der Beiträge in Hinblick auf das Reflexionsniveau der Schülerinnen und Schüler ist das Reflexionsmodell nach Brendel (2017), das explizit zur Einordnung der Reflexionsstufen von Jugendlichen im Bereich des Globalen Lernens verwendet werden kann und auf dem Modell von Bain et al. (1999) beruht. Zusätzlich wird in der Studie die Reflexivität der Lehrkräfte mit Hilfe des Stufenmodells von Zimmermann und Welzel (2008), das zur Kategorienbildung bei der Auswertung der dazugehörigen Interviews dient, untersucht. Dies dient dazu, festzustellen, ob diese die Reflexionsfähigkeit der Schülerinnen und Schüler beeinflusst, was sich jedoch nicht bestätigt. In einem nächsten Schritt leitet die Autorin reflexionsförderliche Maßnahmen ab, die als Orientierung für Lehrkräfte für ihren eigenen Unterricht dienen sollen.

Das Stufenmodell, das Brendel aus ihrer Untersuchung ableitet, dient als Grundlage für die Auswertung der Reflexionsstufen der Schülerinnen und Schüler in dieser Arbeit. Brendels Modell wurde explizit für das Bestimmen der Reflexionsfähigkeiten im Rahmen einer BNE entwickelt und eignet sich daher optimal für die Kategorienbildung zur Bestimmung der Reflexionsfähigkeiten in dieser Studie.

Weiterhin unterscheidet sie in ihrer Arbeit zwischen Reflexion und Reflexivität (vgl. Brendel 2017, S. 19). Beiden Begriffen liegen ausführlich vorgestellte Theorien zugrunde, weshalb diese Differenzierung der beiden Reflektierarten auch für diese Studie als Grundlage für die Herleitung der Arbeitsdefinition des Begriffs *Reflektieren* herangezogen wird.

Die reflexionsfördernden Maßnahmen, die Brendel aus ihrer Studie ableitet, finden ebenfalls in dieser Arbeit Berücksichtigung. Brendel (2017) kommt zu dem Ergebnis, dass u.a. die „Betrachtung des Dreiecks/Vierecks der Nachhaltigkeit in seinen Wechselwirkungen und Zielkonflikten“ (ebd., S. 238) als reflexionsfördernde Methode angesehen werden kann. Dies wird bei der in dieser Studie entwickelten Reflektiermethoden für das Nachdenken über bestimmte Inhalte und Themen eingebunden (vgl. Kapitel 3.3.4). Auch der Aspekt des „Herstellen[s] eines Alltags- bzw. Lebensweltbezugs“ (ebd., S. 238) als reflexionsförderliche Maßnahme wird bei der Formulierung der genauen Aufgabenstellung für das Verfassen der Blogbeiträge berücksichtigt. Ebenfalls werden einzelne Trigger, die Brendel zur Förderung der Reflexionsstufen definiert, bei der Auswahl einer geeigneten Reflexionshilfe als Entscheidungskriterium herangezogen. Dazu zählen beispielsweise die Trigger „Strukturierungshilfen geben“, „Andere Perspektiven einnehmen“, „Zukünftige Entwicklung prognostizieren“ (ebd., S. 267) und „Wertvorstellungen diskutieren und positionieren“ (ebd., S. 254).

2.7.2 Weitere Studien zu Reflektierfähigkeiten im Rahmen des Geographieunterrichts und des Globalen Lernens

Aufbauend auf die gerade vorgestellte Studie befassen sich Brendel und Schrüfer 2018 in einer Studie mit der Reflexionsperformanz von Master-Studierenden im Bereich Globalen Lernens. Zur Auswertung der Reflexionsniveaus, die ebenfalls durch das Schreiben von Blogbeiträgen visualisiert werden, wird wie in der Studie zuvor die Methode der skalierenden Strukturierung nach Mayring (2015) gewählt. Die Kategorisierung erfolgt durch Brendels Stufenmodell reflexiven Denkens im Globalen Lernen, außerdem werden weitere Kategorien durch eine thematische Kodierung nach Kuckartz (2012) induktiv-deduktiv entwickelt. Insgesamt wurden 35 Weblogs von Studierenden ausgewertet. Ein Ergebnis der Studie ist die Erkenntnis, dass Studierende unterschiedliche Reflexionsniveaus aufzeigen, die meisten Ergebnisse im unteren Bereich aufweisen und nur einige wenige Studierende die höchste Stufe erreichen. Der Einsatz von Weblogs ermöglicht es, die Reflexionsperformanz der Studierenden sichtbar zu machen, sodass der Lehrende individuell auf die unterschiedlichen Voraussetzungen der Lernenden reagieren kann. Brendel und Schrüfer (2018) leiten daraus ab, dass Reflexionen gezielt angeregt werden sollten, um die Kompetenzentwicklung der Lernenden zu fördern (vgl. ebd., S. 101). Dies kann beispielsweise durch Feedback in Kommentaren erfolgen. Die Erkenntnisse, dass Blogs Reflektiervorgänge initiieren und transparent machen und dass die Förderung der Reflektierfähigkeiten durch gezielte Impulse unterstützt werden sollte, finden in der

hier vorliegenden Studie Beachtung. Zum einen wird zur Erfassung von Reflektiervorgängen ebenfalls das Medium Blog gewählt, zum anderen werden die Blogbeiträge durch die Forscherin kommentiert und entsprechend auf die Reflektierhilfen verwiesen, die zuvor mit den Jugendlichen behandelt werden. Die Studie stützt zudem die Prämisse, dass sich Brendels Stufenmodell reflexiven Denkens im Globalen Lernen in der Forschung als geeignetes Kategoriensystem im Rahmen der qualitativen Inhaltsanalyse nach Mayring (2010) im entsprechenden Kontext bewährt und sich der Einsatz dieses Modells somit auch für weitere Studien im Bereich des Globalen Lernens eignet.

Die beiden Studien dienen als Grundbasis für die Entwicklung der hier vorliegenden Arbeit. Sie zeigen, dass Blogs sinnvolle Medien sind, um Reflektierprozesse zu visualisieren. Außerdem verwenden sie Stufenmodelle zum Auswerten der Reflektierfähigkeiten, die nach ausführlicher Literaturrecherche gewählt wurden und sich über die Studien hinweg als geeignet erweisen. Die Empfehlungen, die Brendel (2017) für die Förderung des Reflektierens gibt, werden ebenfalls in dieser Studie aufgenommen. Während die beiden Arbeiten allgemein die Reflektierfähigkeiten von Lernenden abbilden und daraus Hinweise auf reflektierfördernde Maßnahmen abgeleitet werden, soll durch diese Studie untersucht werden, inwieweit spezielle Reflektierhilfen die Reflektierfähigkeiten – sowohl Reflexions- als auch Reflexivitätsfähigkeit – von Jugendlichen beim Schreiben von Blogbeiträgen beeinflussen bzw. fördern können.

Metz (2016) beschäftigt sich in ihrer Dissertationsstudie mit der Relevanz von Reflektierfähigkeiten im Rahmen einer BNE. Dabei untersucht sie Strukturen des Konsumhandelns 26 Studierender in Form des Konzepts Forschung in der Lehre. Mithilfe reflexiver Fotografien, die studentische Lebensstile aufzeigen, einer Textanalyse und eines Fragebogens wird im Rahmen eines qualitativen Experiments untersucht, welche Motive und Einstellungen die Testpersonen für nachhaltigen Konsum identifizieren und welche Faktoren ihren Einschätzungen zufolge förderlich und hinderlich dafür sind. Dabei werden u.a. Aspekte der Praktikabilität, Spaß, Luxus und günstige Preise genannt. Ein Ergebnis der Studie ist der, dass Nachhaltigkeit häufig mit Umweltbewusstsein assoziiert wird. Die Studierenden weisen ein hohes Faktenwissen bezüglich des Prinzips von Nachhaltigkeit auf und die Reflektierfähigkeiten, die durch die reflexiven Texte erhoben werden, verändern sich durch die Bildanalysen nicht. Für die hier vorliegende Forschungsarbeit ist die Erkenntnis von Bedeutung, dass Nachhaltigkeit meistens nur einperspektivisch aus dem Bereich der Ökologie heraus betrachtet wird. Deshalb ist es wichtig, im Rahmen dieser Studie das Prinzip des Nachhaltigkeitsdreiecks mit den Bereichen Ökologie, Ökonomie und Soziales einzubinden, da dadurch unterschiedliche Perspektiven berücksichtigt werden. Auf den Bereich Politik, der aus dem Konzept das Nachhaltigkeitsviereck formt, wird verzichtet, da das Reflektieren selbst bereits ein komplexer Prozess ist und das Konzept lediglich als Orientierung dafür dienen und nicht zu Verständnisschwierigkeiten führen soll.

Neben den drei erwähnten Arbeiten zu Reflektierfähigkeiten, gibt es zwei weitere Arbeiten, die sich explizit mit dem Reflektieren innerhalb des Faches Geographie beschäftigen. Brendel gibt in ihrer Studie einen guten Überblick über die Arbeiten von Schneider (2013) und Gryl (2012), weshalb an dieser Stelle nur kurz die wichtigsten Erkenntnisse aus den Studien vorgestellt werden. Schneider schlägt in ihrer Arbeit einen konzeptionellen Entwurf einer „geographiedidaktischen Reflexivität“ (ebd., S. 230) vor. Sie sieht Reflexivität als eine Erweiterung der Reflexion und sieht den Geographielehrenden als einen reflexiven Beobachter, der unterschiedliche Sichtweisen auf geographische Themen einnimmt. Die Studie fokussiert sich folglich auf Lehrende und sieht die Reflexivität als wichtige Fähigkeit dieser. Auch Gryl (2012) befasst sich in ihrer Studie mit dem Reflektieren von Geographielehrkräften und Lehramtsstudierenden. Sie untersucht, welche Ausprägungen die Testpersonen hinsichtlich einer reflexiven Geomedialekompetenz aufweisen, wodurch der Kompetenzerwerb initiiert wird und welche Möglichkeiten es zur Vermittlung der entsprechenden Kompetenz gibt. Neben der Erstellung einer Typologie der Lehrkräfte gibt die Autorin Hinweise für die Förderung der reflexiven Geomedialekompetenz bei Geographielehrenden. Außerdem stuft sie die Reflexivitätsfähigkeit, die sie mithilfe von Interviews misst, in fünf Kategorien ein. Beide Autorinnen richten ihren Fokus auf die Reflektierfähigkeiten von Lehrenden und verwenden die Begriffe *Reflexion* und *Reflexivität* mit unterschiedlichen Verständnissen. Dies zeigt, dass es wichtig ist, genau zu definieren, was unter den verschiedenen Reflektierarten verstanden wird und auf welcher Basis dieses Verständnis aufbaut. Da Gryl ähnlich wie Brendel gezeigte Reflektierfähigkeiten in Kategorien einordnet, wurde dieses Verfahren schon mehrfach angewandt und wird für diese Studie als Auswertungsgrundlage gesehen.

2.7.2.1 Forschungsarbeiten zur Förderung von Reflektierfähigkeiten

Bisher gibt es nur wenige Forschungsarbeiten, die sich speziell mit Reflektierfähigkeiten von Schülerinnen und Schülern im Bereich der BNE beschäftigen. Auch zur allgemeinen Förderung von Reflexions- und Reflexivitätsfähigkeiten durch explizite Reflektiermethoden sind Forschungsarbeiten bislang eher ein Desiderat. Im englischsprachigen Raum finden sich zwar vermehrt wissenschaftliche Artikel zum Reflektieren, in denen u.a. auch Methoden vorgestellt werden, die als reflektierfördernd eingestuft werden. Theoretische Grundlagen und empirische Daten, die die vorgestellten Einschätzungen belegen, fehlen hier jedoch überwiegend. Für diese Studie werden de Bonos Hütemethode und Gibbs Reflexivitätszyklus als Reflektierhilfen gewählt, deren Potenziale genauer untersucht werden. Es folgt ein Überblick über bisherige Forschungsergebnisse zu diesen beiden Methoden.

Im deutschsprachigen Raum gibt es nach aktuellem Stand der Forscherin keine Arbeiten über die Auswirkungen des Einsatzes der Hütemethode von de Bono auf das Reflektieren oder auf

dazugehörige Fähigkeiten. International gibt es ebenfalls kaum Arbeiten darüber. Auf zwei Studien soll nachfolgend eingegangen werden.

Karadag et al. (2009) untersuchen im Zeitraum von 2006 bis 2007 den Nutzen der kreativen Lehrmethode für die Entwicklung der Fähigkeiten des kritischen Denkens bei Krankenpflegeschülern an der Gaziosmanpasa Universität in der Türkei. In ihrer experimentellen Studie werden 41 Schülerinnen und Schüler nach ihrer Meinung zur Hütemethode befragt, nachdem sie sie im Unterricht kennengelernt und in einem praktischen Teil angewendet haben. Die Probandinnen und Probanden füllen dafür einen Fragebogen mit 20 Items aus, der vom Forscherteam entwickelt wurde und auf einer Likert Skala basiert. Die Antworten werden prozentual und in einem Qui-2-Unabhängigkeitstest analysiert. Die Mehrheit der Testpersonen gibt u.a. an, dass sie durch die Hütemethode leichter Mitgefühl für die Patienten entwickeln, unterschiedliche Ideen und Meinungen berücksichtigen, den Patienten holistisch betrachten, auf positive und negative Aspekte eines Ereignisses eingehen und ihr Denksystem weiterentwickeln können. Karadag et al. (2009) folgern aus ihrer Studie, dass die Hütemethode von de Bono eine gewinnbringende Methode ist, um Multiperspektivität und Kreativität im Unterricht zu fördern und Gedanken frei aber dennoch in einer geordneten Form präsentieren zu können. In der Studie von Eldeen und Maher (2016) wird untersucht, welchen Effekt die Hütemethode nach de Bono auf die Entwicklung kreativen Denkens und auf die schulische Leistung von Schülerinnen und Schülern im unterrichtlichen Kontext zu Gesundheits- und Fitnesskursen hat. Dazu werden 76 Testpersonen in zwei Gruppen aufgeteilt, wobei die Experimentalgruppe de Bonos Hütemethode erhält und die Kontrollgruppe klassischen Unterricht ohne Einbindung der Methode erfährt. Vor und nach dem zehnwöchigen Unterricht werden die Daten mithilfe des Torrance Tests für kreatives Denken erfasst und miteinander verglichen. Das Instrumentarium weist verbale und nonverbale Bestandteile auf und beinhaltet Aspekte wie Flexibilität und Originalität, die Teilaspekte von Kreativität sind. Um die schulische Leistung erfassen zu können, bereiten die Forscher Leistungstests aus den Themenbereichen Gesundheit und Fitness vor. Die statistische Analyse wird in SPSS durchgeführt. Als Ergebnis kann festgehalten werden, dass die Experimentalgruppe sowohl im Torrance Test als auch in den Leistungstests deutlichere Steigerungen in den Ergebnissen erzielt als die Kontrollgruppe. Das Forscherteam leitet aus den Studienergebnissen ab, dass die Verwendung der Hütemethode von de Bono über zehn Wochen hinweg einen förderlichen Einfluss auf die Entwicklung kreativen Denkens und den akademischen Leistungen der Probandinnen und Probanden hat.

Die bisher veröffentlichten Studien über die Potenziale der Hütemethode von de Bono untersuchen Reflektieren – wenn überhaupt – nur als Teilkompetenz kritischen Denkens oder die Wirkung auf andere Fähigkeiten wie die schulischen Leistungen oder kreatives Denken. Beide Studien wurden im medizinischen Kontext durchgeführt. Die Ergebnisse zeigen, dass die Hü-

temethode durchaus positiven Einfluss auf Fähigkeitsentwicklungen nimmt. Jedoch bleibt unklar, inwieweit die Methode dazu beitragen kann, Reflektierfähigkeiten speziell fördern zu können. Aus der Literaturrecherche heraus kann die Hypothese generiert werden, dass die Bonos Hütemethode eine sinnvolle Methode zur Entwicklung von Fähigkeiten ist und folglich auch das schriftliche Reflektieren und insbesondere den Teilbereich der Reflexion bei Schülerinnen und Schülern fördern kann. Dies muss jedoch erst noch empirisch untersucht werden, da der Einfluss speziell auf die Fähigkeit des Reflektierens bislang ein Forschungsdesiderat darstellt. Die hier vorliegende Arbeit möchte dem entgegenwirken.

Auch zum Potenzial von Gibbs Reflexivitätszyklus für die Förderung von Reflektierfähigkeiten gibt es im deutschsprachigen Raum bislang keine Forschungsarbeiten. International betrachtet existieren nur wenige Arbeiten, die sich auf die Folgen des Einsatzes von Gibbs Reflexivitätszyklus konzentrieren.

So untersuchen Fakude und Bruce (2003) im Rahmen einer quasi-experimentellen Studie den Einfluss von geleiteten Tagebucheinträgen auf die Fähigkeit des kritischen Denkens von Testpersonen während ihrer klinischen Ausbildung in der Krankenpflege. Die 43 Teilnehmenden werden dafür in zwei Gruppen eingeteilt. Während die Experimentalgruppe über einen Zeitraum von acht Wochen Tagebucheinträge über ihre klinischen Erfahrungen verfasst, bei denen sie sich an Gibbs Reflexivitätszyklus orientieren sollen, schreibt die Kontrollgruppe keine Einträge. Im Anschluss daran erhalten beide Gruppen die Aufgabe, eine klinische Situation schriftlich zu reflektieren. Das Material wird in Anlehnung an Gibbs sechs Reflexivitätsphasen in sechs Kategorien eingeteilt und die Fähigkeiten in den einzelnen Kategorien genauer analysiert. Die statistisch ermittelten Ergebnisse zeigen, dass die Testpersonen der Experimentalgruppe bessere Reflektierfähigkeiten in den Bereichen der Überlegung von Handlungsalternativen und der Formulierung von Antworten in künftigen, ähnlichen Situationen zeigen. In den übrigen vier Bereichen können keine signifikanten Unterschiede zwischen der Kontroll- und Experimentalgruppe festgestellt werden. Die Autoren formulieren aus den Ergebnissen der Studie die Empfehlung, Lernenden Unterstützung und Richtlinien an die Hand zu geben, um über Pflegepraktiken bedeutungsvoll reflektieren zu können.

In einer weiteren Studie, die sich mit der Frage beschäftigt, inwieweit reflexives Schreiben Einfluss auf die Selbstreflexionskompetenz²¹ von auszubildenden Sozialarbeitern nimmt, stellen Karmen und Marju (2015) als eines der Ergebnisse fest, dass Gibbs Reflexivitätszyklus als sinnvolles und reflektierförderndes Modell unter den Teilnehmenden gesehen wird. Den Testpersonen werden unterschiedliche Hilfen für das Reflektieren angeboten, darunter auch Gibbs Reflexivitätszyklus. Im Anschluss an das reflexive Schreiben über ihre Erfahrungen während der Ausbildung, geben die Probandinnen und Probanden ihre Selbsteinschätzung über das

²¹ gemeint ist dabei der Teilbereich der Reflexivität

Reflektieren in schriftlicher Form ab. 16 von 39 Testpersonen bewerten Gibbs Reflexivitätszyklus als nützliches Instrument für die Selbstreflexion und ziehen es anderen Modellen, wie Bortons Reflexionsmodell (1970) oder John's Modell für strukturiertes Reflektieren (1998) für das Verfassen ihrer reflexiven Texte vor.

Auch die Studie von Chong (2009), die sich mit der Frage beschäftigt, ob reflexive Praxis eine sinnvolle Methode für Auszubildende in der Krankenpflege ist, bestätigt, dass Gibbs beim Reflektieren eine zielführende Hilfestellung bietet. 98 Teilnehmende führen in dieser Studie während ihrer Ausbildung schriftliche Reflexionen über ihre praktischen Erfahrungen durch und orientieren sich dabei an Gibbs Reflexivitätszyklus. Im Anschluss daran beantworten sie einen Fragebogen, der u.a. Items zur eigenen Wahrnehmung reflexiver Praxis enthält. Die Ergebnisse zeigen, dass reflexive Praxis und folglich auch Gibbs Reflexivitätszyklus insgesamt als sinnvoll eingestuft werden, auch von Testpersonen, die dem gegenüber erst skeptisch waren. Die Autorin folgert aus ihrer Studie, dass reflexive Praxis mehr im Unterricht – hier im Bereich der Krankenpflegeausbildung – aufgenommen werden solle.

Die erst kürzlich veröffentlichte Studie von Tawanwongsri und Phenwan (2019) hat Gibbs Reflexivitätszyklus nicht nur als Untersuchungsgegenstand sondern auch als Auswertungsinstrument. In der Studie werden u.a. die Reflektierfähigkeiten von Medizinstudierenden während eines zehnwöchigen Einführungskurses untersucht, indem die Studierenden über ihre gehaltenen Vorträge reflektieren. Dafür werden sie zunächst in die reflexive Praxis eingeführt und erhalten von Gibbs Reflexivitätszyklus abgeleitete Orientierungsfragen. Die mündlichen Äußerungen werden mithilfe von Gibbs Phasen kategorisiert. Somit kann festgestellt werden, welche maximale Phasen erreicht werden und wie sich die Reflektierfähigkeiten über die zehn Wochen hinweg entwickeln. Als Ergebnis halten die Autoren fest, dass Feedback beim Reflektieren eine wichtige Ergänzung darstellt und diese benötigt wird, um höhere Reflektierfähigkeiten erreichen zu können. Insgesamt hilft Studierenden das Reflektieren über vergangene Ereignisse, um ihre reflexiven Kompetenzen weiter auszubilden. Längere Untersuchungszeiträume sind nach Meinung des Forscherteams sinnvoll, um genauere Aussagen über die Entwicklung von Reflektierfähigkeiten treffen zu können.

Insgesamt wird Gibbs Reflexivitätszyklus von Testpersonen als sinnvoll eingestuft und erzielt positive Ergebnisse hinsichtlich der Entwicklung von Fähigkeiten. Allerdings beziehen sich die dazugehörigen Forschungsarbeiten alle auf den medizinischen Bereich und hier im Rahmen der Ausbildung von Berufseinsteigern und nicht auf den Einsatz im schulischen Unterricht. Ob auch hierbei positive Entwicklungen festgestellt werden können, soll in dieser Studie überprüft werden.

2.7.2.2 Forschungsarbeiten mit Reflektierniveaumodellen zur Bestimmung von Reflektierfähigkeiten

Studien, die Messinstrumente von Reflektierfähigkeiten zum konkreten Untersuchungsgegenstand machen, gibt es nach Wissensstand der Forscherin bisher nicht. Es existiert eine Vielzahl an Niveaumodellen, die Reflektieren nach unterschiedlichen Qualitäten unterscheiden. Wie verlässlich diese Modelle jedoch sind, wurde noch nicht überprüft. Im Folgenden werden die Forschungsarbeiten vorgestellt, aus denen die ursprünglichen Niveaumodelle entwickelt wurden, die für diese Studie von Bedeutung sind und ebenfalls Studien genannt, bei denen die Modelle als Auswertungsinstrument eingesetzt werden. Zunächst wird auf die Studie von Hatton und Smith (1995) eingegangen, anschließend auf die Studie von Bain et al. (1999).

Hatton und Smith (1995) entwickeln durch die Analyse schriftlicher Arbeiten von 60 Lehramtsstudierenden der Universität von Sydney und durch die Sichtung der Fähigkeit des Reflektierens in der Literatur vier Arten von Reflexivität: *Descriptive Writing*, *Descriptive Reflection*, *Dialogic Reflection* und *Critical Reflection* (vgl. Kapitel 2.4.2.1). Die Datenerhebung verläuft trianguliert und beinhaltet die Erhebungsinstrumente schriftlicher Bericht, Selbstevaluation, Videoaufnahmen vom Unterrichten und Interview (vgl. ebd., S. 40). Im Sinne einer argumentativen Validierung nach Döring und Bortz (2016, S. 328) werden vorher festgelegte Texteinheiten der Hausarbeiten von zwei Personen hinsichtlich der vier entwickelten Reflexivitätsarten kodiert und bei unterschiedlichen Einordnungen die Ergebnisse gemeinsam diskutiert. Die Ergebnisse der schriftlichen Berichte zeigen, dass die Lehramtsstudierenden in ihrem letzten Ausbildungsjahr Reflexivitätsprozesse durchführen. Die meisten Arbeiten, die die Autoren mit Hilfe des entwickelten Modells untersuchen, bestehen überwiegend aus deskriptiven Reflektieranteilen und nur acht Testpersonen weisen kritische Reflexivität auf. Zwei Arbeiten besitzen keinerlei reflexive Anteile. Gegen Ende der schriftlichen Texte stellen die Autoren häufiger dialogische Anteile fest, da hier oft über Handlungsalternativen nachgedacht wird. Nur acht Studierende erreichen die Stufe des kritischen Reflektierens. Der Bericht mit den meisten reflexiven Anteilen weist 52 Reflexivitätsanteile auf, der mit den geringsten zwei. Durchschnittlich besitzen die schriftlichen Texte (zwischen acht und zwölf Seiten) 19 reflexive Anteile. Am häufigsten kann die zweite Form der *descriptive reflection* festgestellt werden. Hinsichtlich der Reihenfolge von reflexiven Anteilen innerhalb eines Textes stellt die Forschergruppe fest, dass Studierende oft mit einer deskriptiven Reflexivität beginnen und anschließend zu einer dialogischen Reflexivität übergehen (vgl. Hatton und Smith 1995, S. 41 f.). Mit der eben beschriebenen Studie wurde der Grundstein für die Analyse von Reflexivitätsfähigkeiten innerhalb schriftlicher Arbeiten gelegt, da die identifizierten Reflexivitätsarten in weiteren aufbauenden Studien häufig in Stufenmodelle zur Bestimmung der Reflexivitätsniveaus umgewandelt werden. Viele Arbeiten, die sich der Fähigkeit des Reflektierens im Sinne eines Hinterfragens von Handlungen widmen, haben als Ausgangspunkt die Unterscheidung der vier Reflexivitätsarten

nach Hatton und Smith (1995). Zu nennen sind hierfür die Arbeiten von Abels (2008), Brendel (2017), Eysel (2006) und Zimmermann und Welzel (2008). Ihnen allen gemeinsam ist, dass sie Stufen- bzw. Niveaumodelle zur Bestimmung der Reflexivitätsfähigkeit hinsichtlich unterschiedlicher Qualitäten verwenden und sie alle den Fokus auf die Fähigkeit der *reflective practice* nach Schön (1983) legen. Außerdem untersuchen sie alle die Kompetenzen von pädagogischen Fachkräften wie Lehrkräften oder Erzieherinnen und Erziehern im beruflichen Kontext. Ihre Daten erheben sie mittels schriftlicher Texte und Interviews und die Auswertung erfolgt durch qualitative Inhaltsanalysen. Das ähnliche Vorgehen dieser Forschungsarbeiten und die Verwendung von Reflexivitätsstufen für die Bestimmung von Reflexivitätsfähigkeiten dienen als Grundlage für die Entwicklung des Forschungsdesigns im Rahmen dieser Arbeit. Auch in dieser Studie sollen schriftliche Texte mittels eines Niveaumodells ausgewertet werden, um Aussagen über die Entwicklung von Reflexivitätsfähigkeiten über einen längeren Zeitraum und durch Erhalt von Reflektierhilfen treffen zu können.

Die Studie von Bain et al. (1999) untersucht den Nutzen von Reflexionsberichten für das Lernen während eines Praktikums von Lehramtsstudierenden. Insgesamt verfassen 35 Studierende wöchentlich einen Bericht im Rahmen des elfwöchigen Praktikums. Zusätzlich werden sie zu unterschiedlichen Zeitpunkten über ihre Lehrvorstellungen und ihre Lehre selbst befragt. Die Teilnehmenden werden in vier Gruppen eingeteilt, die sich hinsichtlich der Interventionsbedingungen unterscheiden. Dies betrifft den Inhalt und den Kontext des Berichterstattens. Für die Auswertung werden der Fokus und das Level der Reflexionen bestimmt. Für die Einteilung in unterschiedliche Reflexionsniveaus entwickeln die Autoren eine Fünf-Punkte-Skala. Das Material wird von zwei Forschern kodiert, im Anschluss vergleichen sie ihre Ergebnisse und diskutieren unterschiedliche Einteilungen. Für die Ergebnisdarstellung orientieren sich die Autoren an den Mittelwerten der vorhandenen Stufen pro Bericht. Das Ergebnis zeigt, dass es keine Unterschiede bei der Qualität der Reflexion in Abhängigkeit des Inhalts und des Kontexts der Berichtserstattungen gibt. Die Studie von Bain et al. (1999) und insbesondere die fünf daraus entwickelten Reflexionsstufen finden in darauffolgenden Studien Anwendung. Chen et al. (2009) setzen in ihrer quantitativen Studie die Stufenmodelle ein und gewichten die Qualitäten zusätzlich mit Scorewerten. Sie untersuchen die Wirkung von Reflexionsimpulsen auf die Reflexionsfähigkeiten von 157 Studierenden. Dazu werden die Testpersonen zufällig in eines von 14 Lernszenarien eingeteilt und verfassen Reflexionsberichte. Für die Auswertung werden zunächst die Reflexionsstufen der Texte bestimmt und anschließend mithilfe der Scorewerte Varianzanalysen durchgeführt. Als Ergebnis halten die Autoren fest, dass anspruchsvolle Reflexionsimpulse den höchsten Effekt auf die Reflexionsfähigkeit aufweisen, während Peerfeedback keinen Einfluss zeigt.

Das Vorgehen, die Reflexionsstufen zu quantifizieren, um statistische Aussagen über die Wirkung bestimmter Interventionen treffen zu können, ist eine Methodik, die sich auch für diese

Studie anbietet. Die Studie von Chen et al. (2009) ist folglich für die Überlegungen der Datenauswertung von Bedeutung. Das Vorgehen und das Stufenmodell von Bain et al. (1999), an dem sich auch Brendel (2017) orientiert, sind ebenfalls wichtige Kriterien für die Datenerhebung und -auswertung in dieser Studie.

Der vorgestellte Forschungsstand hinsichtlich der Förderung und Messung von Reflektierfähigkeiten im Aus- und Bildungskontext fasst die Arbeiten zusammen, die für die Entwicklung dieser Studie von Relevanz sind. Es gibt bereits Hinweise darauf, inwieweit die Reflektierfähigkeiten von Schülerinnen und Schülern im Rahmen einer BNE gefördert werden können und auch, dass sich die beiden ausgewählten Reflektiermethoden positiv auf die generelle Fähigkeitsentwicklung in unterschiedlichen Kontexten auswirken. Ob diese beiden Feststellungen auch in Kombination gelingen, soll mit dieser Studie herausgefunden werden. Die Modelle für die Einordnung in unterschiedliche Reflektierniveaus werden ebenfalls aus dem bisherigen Forschungsstand abgeleitet. Die hier vorliegende Forschungsarbeit möchte neue Erkenntnisse im Bereich der expliziten Reflektierförderung von Schülerinnen und Schülern im Kontext einer BNE liefern und somit eine Lücke innerhalb des noch wenig erforschten Gebiets füllen, die Brendel (2017, S. 277) abschließend in ihrer Dissertationsarbeit formuliert.

3 Empirische Untersuchung

Im empirischen Teil wird zunächst auf die Forschungsfrage und die Teilforschungsfragen eingegangen. Anschließend werden das übergeordnete Forschungskonzept, das Forschungsdesign, das Erhebungs- und Auswertungsverfahren genauer beschrieben. Vor der Beantwortung der Forschungsfrage und ihren Unterfragen werden die Ergebnisse der Auswertungen, bestehend aus einem qualitativen und einem quantitativen Teil, zusammengefasst. Die Diskussion mit Erkenntnissen und Grenzen dieser Studie rundet die empirische Untersuchung ab.

3.1 Forschungsfragen

Ausgehend von der Wichtigkeit von Reflektierprozessen im schulischen Kontext (vgl. Kapitel 2.2 und 2.3) und dem Forschungsdesiderat zur Förderung der Reflexions- und Reflexivitätsfähigkeiten (vgl. Kapitel 2.7), geht diese Studie der Frage nach, inwieweit gezielte Reflektiermethoden die Fähigkeiten im Reflektieren – sowohl im Bereich der Reflexion als auch der Reflexivität – beeinflussen. Besonderes Augenmerk wird hierbei auf die Potenziale der einzelnen, im Unterricht vorgestellten Methoden auf die Reflexions- und Reflexivitätsfähigkeit gelegt und auf den Vergleich der Reflektierfähigkeiten der Jugendlichen vor und nach dem Erhalt beider Reflektierhilfen. Eine Forschungsfrage sollte drei Kriterien erfüllen, damit sie als legitimierte Forschungsfrage zählt: Zum einen sollte sie im Sinne des soziologischen Interesses relevant in Hinblick auf das Weltverständnis sein. Zum anderen sollte sie zu einem Erkenntnisgewinn führen und drittens sollte sie die Forschung im entsprechenden Bereich voranbringen und Lücken schließen (vgl. King und Kitchener 1994, S. 14 ff.). Die drei Kriterien werden durch die folgende Forschungsfrage erfüllt. Weiterhin ist es wichtig, dass sowohl die Hauptforschungsfrage als auch die Unterfragen präzise formuliert und nicht zu allgemein gehalten sind, sodass deren Beantwortung während der entsprechenden Forschung auch gewährleistet wird (vgl. Schwarzer 2001, S. 149). Die Konkretisierung der Frage ist nötig, um den Überblick über das Untersuchungsfeld zu behalten und den daraus abgeleiteten Untersuchungsgegenstand sinnvoll zu strukturieren (vgl. Flick 2017, S. 259). Nachdem die Richtlinien kurz umrissen wurden, an denen sich die Autorin für die Entwicklung der Forschungsfragen orientiert hat, werden nun die zentrale Forschungsfrage und deren Teilfragen, die bei der Beantwortung der Hauptforschungsfrage helfen, vorgestellt:

Hauptforschungsfrage:

Welche Potenziale bieten ausgewählte Reflektiermethoden für die gezielte Förderung von Reflektierfähigkeiten bestehend aus Reflexions- und Reflexivitätsfähigkeiten von Jugendlichen beim Bloggen im Rahmen einer Bildung für nachhaltige Entwicklung?

Teilforschungsfragen:

1. Welches Potenzial bieten die beiden Reflektiermethoden insgesamt für die Entwicklung der Reflektierfähigkeiten beim Bloggen?
2. Welches Potenzial bietet die Reflektiermethoden in Form von Gibbs Reflexivitätszyklus für die Entwicklung der Reflektierfähigkeiten?
3. Welches Potenzial bietet die Reflektiermethoden in Form der Hütemethode für die Entwicklung der Reflektierfähigkeiten?
4. Inwieweit ändert sich das Begriffsverständnis von Reflektieren durch das Blogprojekt?
5. Lässt sich ein Zusammenhang zwischen der Reihenfolge des Erlernens der Reflektiermethoden und der Veränderung der Reflektierfähigkeiten insgesamt feststellen?
6. Unterscheidet sich das Potenzial beider Reflektiermethoden insgesamt hinsichtlich des Geschlechts? Wenn ja, wodurch?
7. Besteht ein Zusammenhang zwischen der eigenen Bewertung der Reflektiermethoden und der Veränderung der Reflektierfähigkeiten durch diese?
8. Inwieweit verändern sich persönliche Einstellungen gegenüber dem Reflektieren durch das Blogprojekt?

3.2 Übergeordnetes Forschungskonzept und Forschungsdesign

In der empirischen Sozialforschung wird traditionell zwischen quantitativen und qualitativen Forschungsmethoden unterschieden. Obwohl sich die beiden Forschungsrichtungen zwar in einigen Aspekten unterscheiden, können sie auch miteinander kombiniert werden, da sie sich nicht zwingend gegenseitig ausschließen (vgl. Wolf 1995, S. 318). Lamnek und Krell (2016, S. 254 ff.) sehen den Hauptunterschied der beiden Forschungsmethoden im Grundverständnis: Während in der quantitativen Forschung eher objektbezogen vorgegangen wird, mit dem Ziel, Ursache-Wirkungszusammenhänge aufzudecken, wird in der qualitativen Forschung der Fokus auf das Subjekt gelegt. Bei ersterem geht es voranging darum, Hypothesen zu überprüfen, wofür große Datensätze benötigt werden (vgl. Brüsemeister 2008, S. 19). Bei der qualitativen Forschung wird mit einer wesentlich kleineren Fallzahl gearbeitet, da hierbei „die Entdeckung

(Generierung) von Theorieaussagen anhand empirischer Daten“ (ebd.) im Mittelpunkt steht und „der Hauptuntersuchungsgegenstand [...] immer das menschliche Subjekt“ (Röbken und Wetzel 2016, S. 14) selbst ist. Qualitative Methoden sind in der Regel explorativ und hypothesengenerierend ausgerichtet, d.h. durch „generative[] Fragen“ (Flick 1996, S. 69) sollen Theorien erst entwickelt werden. Beide Methoden bieten gewisse Vor- und Nachteile, die Winter (2000) übersichtlich zusammenstellt. An dieser Stelle sei erwähnt, dass der qualitativen Forschung oft unterstellt wird, ihr mangle es an Objektivität, da die Interpretationen zu subjektiv und nicht kontrollierbar wären (vgl. Saldern 1991). Dem kann jedoch durch geeignete Auswertungsschritte entgegengewirkt werden (vgl. Kapitel 3.4.2).

In der hier vorliegenden Studie, die sich am *mixed-methods*-Ansatz orientiert, sind sowohl qualitative als auch quantitative Anteile in der Datenerhebung und -auswertung vorhanden. Das Studiendesign besteht sowohl aus einem hypothesentestenden, als auch hypothesengenerierenden Teil. Im Folgenden soll die Wahl des gewählten Forschungskonzepts und der Methodenkombination theoretisch hergeleitet werden, bevor es anschließend ausführlich vorgestellt wird.

Mixed-methods-Studien orientieren sich sowohl an qualitativen als auch quantitativen Ansätzen in einem Forschungsprojekt. Die ursprüngliche Systematisierung von *mixed-methods*-Studien nach Morse (1991) konzentriert sich dabei nur auf die Phase der Datenerhebung. Hussy et al. (2010) merken jedoch an, dass auch die Phasen der Auswertung und des Designs berücksichtigt werden können: „Eine Methodenkombination muss sich nicht auf die Datenerhebung beschränken. Kombinationen sind auch bei der Auswertung und auf der Designebene möglich“ (ebd., S. 286). So können wie in der Erhebungsphase auch in der Auswertungsphase qualitative und quantitative Verfahren kombiniert angewendet oder ein Design gewählt werden, dass sowohl hypothesengenerierend als auch hypothesentestend ausgelegt ist. Beim erweiterten *mixed-methods*-Studien-Verständnis können den eingesetzten Methoden unterschiedliche Gewichtungen gegeben werden und sie können nacheinander oder parallel, in einer Untersuchungsphase oder mehreren eingesetzt werden. Insgesamt ergibt sich eine Vielzahl an möglichen *mixed-methods*-Designs, die von unterschiedlichen Autoren typisiert werden (vgl. Mertens 2005; Miles und Huberman 2009; Morgan 1998; Tashakkori und Teddlie 2003). Creswell und Plano Clark (2011, 60 ff.) geben einen Überblick über die verschiedenen Zuordnungen von *mixed-methods*-Designs.

Für diese Studie wird sich an den unterschiedlichen Ausprägungen von Tashakkori und Teddlie (2003; 1998) orientiert, auf die nachfolgend genauer eingegangen werden soll. Die beiden Autoren prägen den Begriff *mixed-model-designs*, der eine besondere Form von *mixed-methods*-Studien bezeichnet. Dabei werden qualitative und quantitative Elemente über die drei Untersuchungsphasen Designtyp, Datenerhebung und Datenauswertung hinweg kombiniert.

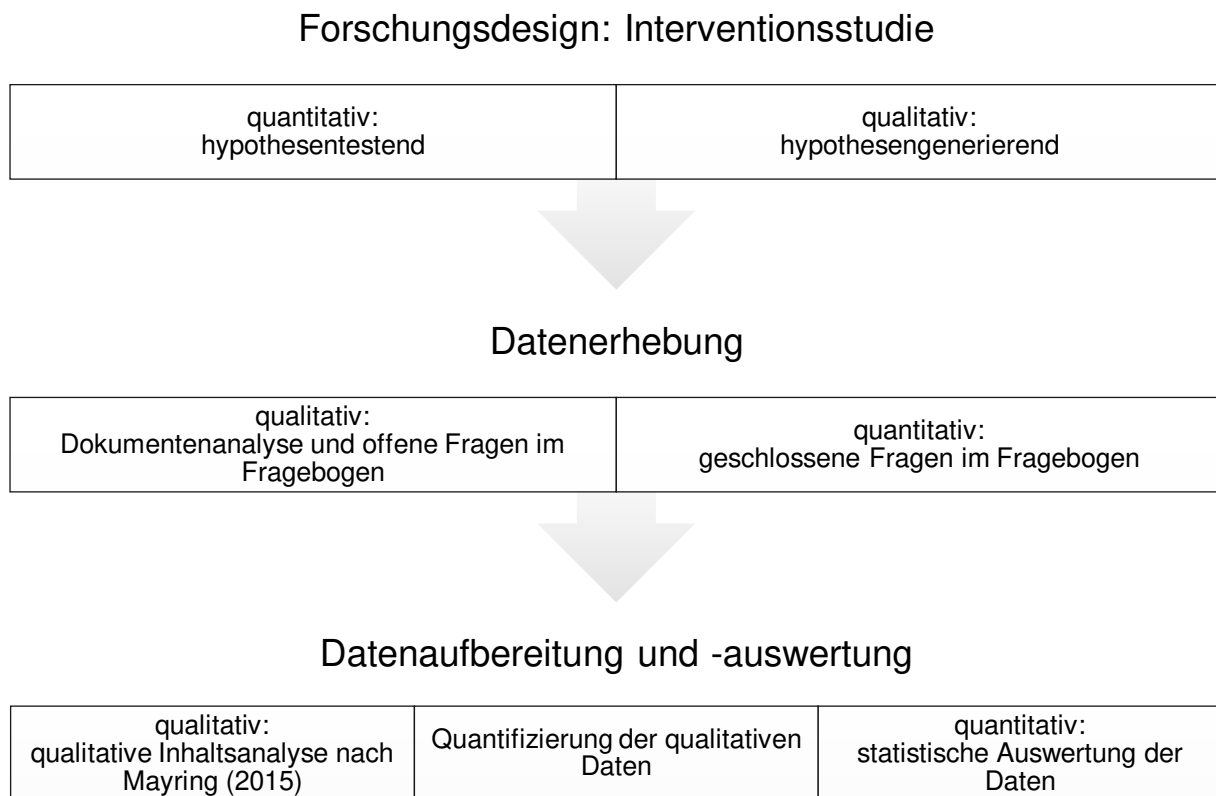
Insgesamt unterscheiden sie sechs solcher Mischdesigns: hypothesentestend oder hypothesengenerierendes Design mit qualitativen oder quantitativen Daten und einer statistischen oder qualitativen Auswertung (vgl. Tashakkori und Teddlie 1998). In einer weiteren, später publizierten Klassifizierung differenzieren die Autoren die zwei unterschiedlichen Designs *mixed-methods* und *mixed-model* in weitere Unterkategorien (vgl. Tashakkori und Teddlie 2003). Dabei unterscheiden sie – eine Ebene höher als bei der ersten Klassifizierung – zunächst wie bereits erwähnt zwischen *mixed-methods-designs* und *mixed-model-designs*, wobei sich ersteres nur auf die Kombination qualitativer und quantitativer Methoden in der Phase der Datenerhebung bezieht und letzteres auch auf das Design der Studie und die Auswertungsphase. In beiden Teilbereichen existieren jeweils *concurrent*, *sequential* und *conversion designs*, d.h. die gleichzeitige, nacheinander aufgestellte und umwandelnde Vermischung qualitativer und quantitativer Methoden (vgl. Tashakkori und Teddlie 2003, S. 687). Nach dieser Klassifizierung liegt bei der hier vorliegenden Forschungsarbeit ein *conversion mixed model design* vor, in dem als hypothesentestend angelegte Studie qualitative Daten erhoben werden, die mit der qualitativen Inhaltsanalyse ausgewertet und anschließend quantifiziert werden und die aufgestellten Hypothesen anhand einer statistischen Untersuchung überprüft werden. Der hier gewählte *mixed-methods*-Ansatz bezieht sich auf alle drei Untersuchungsphasen. Neben dem quantitativen Hauptstudiendesign sollen außerdem anhand der vorliegenden Daten neue Hypothesen generiert werden, sodass das Forschungsdesign auch zu einem Teil qualitativ ausgerichtet ist.

Der Grund für die Wahl einer *mixed-methods* Studie ist folgender: Um Reflektierfähigkeiten erfassen zu können, bietet es sich an, schriftliche Portfolios verfassen zu lassen. Denn Reflektieren ist ein Prozess, der Zeit benötigt. Beim Schreiben kann man seine Gedankengänge klar strukturieren und niederschreiben (vgl. Bräuer 2014, S. 76 ff.). Ein Fragebogen mit geschlossenen Antworten bietet sich hingegen weniger an, da Reflektieren eine Fähigkeit ist, der neben Zeit auch Raum gegeben werden muss, um sich entfalten zu können. Definiert man Antworten vor, werden die tatsächlichen Reflektierfähigkeiten nur schwer erfasst. Die meisten Studien, die sich Fähigkeiten wie dem kritischen Denken und dem Reflektieren widmen, verwenden Interviews oder Texte für die Datenerhebung (vgl. Brendel 2017; Oner und Adadan 2011). Das ist auch hier der Fall. Die Daten werden anschließend mit Hilfe der qualitativen Inhaltsanalyse nach Mayring (2015) kodiert und zunächst qualitativ ausgewertet. Anschließend werden sie für eine weitere quantitative Auswertung in dafür benötigte quantitative Daten umgewandelt. Die Transformation ist dadurch begründet, dass die Studie überwiegend aus der Theorie hergeleitete Hypothesen überprüfen möchte und Aussagen darüber getroffen werden sollen, welche Wirkung die Intervention auf die Reflektierfähigkeiten hat. Quantitative Daten ermöglichen direkte Vergleiche im Rahmen dieser Interventionsstudie, die auf qualitativen Grunddaten be-

ruht. Die Kombination der qualitativen und quantitativen Methoden hat das Ziel, einen modernen Ansatz als Verbindungsbrücke zwischen den beiden Traditionen der qualitativen und quantitativen Forschung zu bilden, um aus den Vorteilen beider Bereiche für diese Studie zu profitieren.

Natürlich gibt es auch Kritik hinsichtlich des *mixed-methods*-Ansatzes, auf die im Folgenden kurz eingegangen wird. McMillan und Schumacher (2006, S. 401) stellen drei Nachteile dieses Ansatzes vor: Zunächst setzt die Durchführung einer *mixed-methods*-Studie eine hohe fachliche Kompetenz des Forschers voraus, da sowohl qualitative als auch quantitative Methoden verwendet werden und der Forscher über ausreichend Expertise verfügen muss, um diese adäquat anwenden zu können. Außerdem muss eine umfangreiche Datensammlung durchgeführt werden und letztendlich besteht das Risiko, dass *mixed-methods* nur oberflächlich verwendet werden (vgl. ebd.). Wichtig ist es, die drei vorgestellten Aspekte bei der Planung und Durchführung einer Studie zu berücksichtigen. Für diese Studie hat sich die Forscherin deshalb intensiv mit qualitativen als auch quantitativen Forschungsmethoden auseinandergesetzt und gemeinsam im Team über mögliche Kombinationen diskutiert. Da die Daten größtenteils qualitativ erhoben und kodiert werden, wurde eine Stichprobe gewählt, die von ihrer Größe her noch qualitativ untersucht werden kann und gleichzeitig groß genug ist, um Aussagen über Effekte im Rahmen dieser Interventionsstudie treffen zu können. Maxwell und Loomis (2003, S. 268 f.) betonen die Wichtigkeit, bei der Verwendung eines *mixed-methods*-Designs die gewählten qualitativen und quantitativen Anteile innerhalb der einzelnen Untersuchungsphasen einer Studie genau zu charakterisieren. Dieser Empfehlung folgend, sollen nun die einzelnen methodischen Überlegungen für das Design vorgestellt und die wichtigsten Aspekte der Datenerhebung und -auswertung erwähnt werden, die sich am *mixed-methods-design* orientieren (vgl. Abbildung 2). Die beiden letzten Punkte werden im nächsten Kapitel noch ausführlicher aufgegriffen.

Abbildung 2 Überblick über übergeordnetes Forschungskonzept für diese Studie



Wie bereits erwähnt, orientiert sich die hier vorliegende Studie am Ansatz des *conversion mixed model designs* nach Tashakkori und Teddlie (2003). Die Kombination qualitativer und quantitativer Methoden findet dabei in allen drei Untersuchungsphasen statt. Das Forschungsdesign entspricht vom Grundprinzip her einer Interventionsstudie bestehend aus zwei Teilinterventionen. Neben einer Kontrollgruppe, die keine Intervention erhält, kann die Experimentalgruppe hinsichtlich der Reihenfolge der erhaltenen Teilinterventionen in zwei Gruppen unterteilt werden. Zur Beantwortung der Forschungsfrage wird im quantitativen Auswertungsteil zunächst hypothesentestend vorgegangen. Zum Thema Reflektieren gibt es Erkenntnisse darüber, dass Reflektier- bzw. Strukturierungshilfen die Reflektierfähigkeiten fördern können (vgl. Kapitel 2.5.1). Ebenfalls gibt es – wenn auch nur wenige – Erkenntnisse zu den fördernden Eigenschaften der Hütemethode und des Reflexivitätszyklus von Gibbs (vgl. Kapitel 2.7.2.1). Darauf aufbauend wird die Haupthypothese definiert, dass die Intervention einen positiven Effekt auf die Reflexions- und Reflexivitätsfähigkeit und folglich auf die Reflektierfähigkeit insgesamt hat. Neben der Haupthypothese werden zwei weitere Hypothesen durch die Studie überprüft, die sich auf die Wirkung der beiden Teilinterventionen beziehen. Dieses Vorgehen entspricht einem traditionellen quantitativen Forschungsdesign. In einem zweiten Schritt werden aus dem Datenmaterial heraus neue Hypothesen generiert, das mit einem klassischen qualitativen Forschungsdesign übereinstimmt. Folglich werden bereits im Forschungsdesign der Studie quantitative und qualitative Aspekte aufgenommen. Die qualitative Inhaltsanalyse nach

Mayring (2015) ist eher hypothesengenerierend ausgelegt und dient als Vorbereitung der quantitativen Daten und als zusätzliche Quelle für die Beantwortung der (Teil-)Forschungsfrage(n).

Auch bei der Datenerhebung finden sowohl qualitative als auch quantitative Methoden Anwendung. Für die Erfassung der Hauptdaten werden schriftliche Texte durch die Testpersonen generiert und folglich die qualitative Methode der Dokumentenanalyse verwendet. Zusätzlich dazu werden in einem vor- und nachgelagerten Fragebogen quantitative und qualitative Daten durch geschlossene und offene Antwortformate erzeugt. Die gewonnenen Daten aus der Dokumentenanalyse werden mithilfe der qualitativen Inhaltsanalyse nach Mayring (2015) in Form der skalierenden Strukturierung mit der Software MAXQDA aufbereitet. Das Vorgehen ist dabei induktiv, d.h. die Kategorien sind bereits vor der Analyse des Datenmaterials aufgestellt und definiert. Für die Kategorienbildung werden zwei Stufenmodelle verwendet, die bereits in anderen Studien zum Einsatz kommen. In einer Vorstudie werden die übernommenen Kategorien entsprechend für diese Studie angepasst. In einem nächsten Schritt werden die Daten aus der qualitativen Inhaltsanalyse in quantitative Daten transformiert. Dies geschieht zum einen durch die Bildung von Scores, zum anderen durch die Quantifizierung der erreichten Häufigkeiten der einzelnen Stufen. Die quantitativen Daten werden im Anschluss mit der Software SPSS in Hinblick auf Signifikanzen und Effektstärken ausgewertet. Zusammenfassend besteht die Datenaufbereitung und -auswertung wie auch die Datenerhebung aus quantitativen und qualitativen Elementen und die Forschungsfrage wird mithilfe der qualitativen und quantitativen Erkenntnisse beantwortet.

3.3 Erhebungsverfahren

Nachdem das übergeordnete Konzept dieser Forschung theoretisch hergeleitet und die einzelnen Bestandteile davon vorgestellt wurden, wird nun ausführlich auf die einzelnen Phasen eingegangen. Zunächst wird das Erhebungsverfahren in Form von Blogbeiträgen genauer erläutert, danach wird auf den Fragebogen und das Sampling eingegangen.

3.3.1 Begründung für die Auswahl von Blogs als Haupterhebungsinstrument

Reflektieren ist ein individueller Prozess, der bei jedem unterschiedlich abläuft und von vielen oftmals nicht direkt wahrgenommen wird (vgl. Bade et al. 2018, S. 40). Da es sich dabei um einen kognitiven Vorgang handelt, den man selbst nur schwer beschreiben kann bzw. dessen Ausprägungen sich eher in der Performanz als in persönlichen Einschätzungen zeigen, lassen sich Reflektierniveaus nur schwer durch standardisierte Fragebögen oder Interviews erfassen. Deshalb wurde sich in dieser Studie gegen dieses Erhebungsverfahren entschieden.

Fragebögen mit geschlossenem Antwortformat werden in der quantitativen Forschung verwendet, um statistische Ergebnisse von einer möglichst großen Stichprobe erzielen zu können. Da der Forschungsschwerpunkt in dieser Studie jedoch auf den individuellen Reflektierfähigkeiten der Untersuchungspersonen liegt und diese nur schwer durch vorformulierte Frage-Antwort-Formate messbar sind, wird für die Erhebung ein qualitatives Erhebungsverfahren gewählt. Fragebogen mit offenem Antwortformat liefern zwar im Gegensatz zu geschlossenen Antwortformaten individuelle Texte, jedoch werden die Reflektierfähigkeiten hierdurch ebenfalls nur oberflächlich sichtbar. Grund dafür ist, dass in einem Fragebogen meist nur knappe Antworten gegeben werden. Bevorzugt werden in dieser Studie jedoch längere, frei geäußerte Aussagen, die genügend Material zur Auswertung und begründeten Einteilung in Reflektierniveaus bieten. Deshalb wurde sich in dieser Studie für das Erheben der Daten durch das Verfassen von Texten in Form von Blogbeiträgen entschieden. Folgende Gründe sprechen für die gewählte Datenerhebung gegenüber der Erhebungsvariante eines gesprochenen Interviews:

- Das Reflektieren benötigt Zeit und kann nicht auf Abruf stattfinden (vgl. Carrington und Selva 2010, S. 42). Blogbeiträge können räumlich und zeitlich unabhängig verfasst werden, sodass der Reflektierprozess ungestört stattfinden kann. Bei der Durchführung von Interviews wird eine bestimmte Zeit vorgegeben, unabhängig von der Verfassung der Untersuchungsperson.
- Um Reflektierprozesse sichtbar zu machen, bietet sich das schriftliche Verfassen von Texten an (vgl. Gläser-Zikuda und Hascher 2007). Beim Schreiben hat man ausreichend Zeit, sich korrekt zu artikulieren und kann Gedanken ordnen und strukturieren. Bei mündlichen Äußerungen ist dies nicht gegeben.
- Um den Schreibprozess beim Bloggen zu beginnen, benötigt man keine weitere Person außer der Untersuchungsperson. Für die Durchführung von Interviews wird diese jeweils benötigt. Da es sich bei dieser Studie um eine Studie mit mehreren Erhebungszeitpunkten mit Schülerinnen und Schülern aus unterschiedlichen Schulen in Oberfranken handelt, wäre die Durchführung von Interviews zeitlich und logistisch schier unmöglich.
- Beim Interview besteht die Gefahr der Wiedergabe von Erwarteten. Da man beim Bloggen alleine ist und keiner Person gegenüber sitzt, bei der man möglicherweise das Gefühl hat, sich für seine Gedanken und Verhalten rechtfertigen zu müssen, wird die Gefahr der sozialen Erwünschtheit verringert (vgl. Wittkowski 2013, S. 33).

Eine Alternative zu Blogs wären klassische Portfolios oder Lerntagebücher gewesen. Da diese jedoch nicht digital geführt werden, wäre der Zugang zu ihnen während des Projekts nur

schwer möglich. Da die Schülerinnen und Schüler jedoch regelmäßiges Feedback mit individuellen Tipps zur Anwendung der Reflektierhilfen durch die Forscherin erhalten, erweisen sich Blogs als Sonderform von E-Portfolios als geeigneter gegenüber klassischen Heftführungen.

3.3.2 Fragebogen als weiteres Erhebungsinstrument

Neben den Blogbeiträgen als Erhebungsinstrument werden zusätzliche Daten durch ein weiteres Instrument erhoben. Zu Beginn der Studie erhalten die Testpersonen einen digitalen Fragebogen, der aus mehreren Elementen besteht. Zum einen wird durch den Fragebogen der Username im Blog abgefragt, damit die Antworten des Fragebogens den entsprechenden Autorinnen und Autoren der Blogbeiträge zugeordnet werden können. Weiterhin wird das Geschlecht geklärt, um die Ergebnisse der Studie in Hinblick auf Geschlechterunterschiede untersuchen zu können. Die Abfrage hierüber ist nötig, da die Usernamen der Teilnehmenden keine Aussagen über das Geschlecht treffen und der Forscherin die Namen der Schülerinnen und Schüler nicht bekannt sind. In einem offenen Aufgabenformat wird nach dem individuellen Verständnis zum Begriff *Reflektieren* gefragt.

Weiterer Bestandteil des Fragebogens sind Ratingskalen zur Selbsteinschätzung. Ratingskalen gehören seit ihrer Einführung in den 30er Jahren zu den wichtigsten und weitverbreitetsten Instrumenten sozialwissenschaftlicher Datenerhebung (vgl. Menold und Bogner 2016). Zu beachten gilt bei der Erstellung einer Ratingskala der Variationsbereich und der Differenzierungsgrad, da diese beiden Merkmale die Antwortgebung beeinflussen (vgl. Parducci 1983). Der Variationsbereich wird durch die beiden Skalenpole abgegrenzt, der Differenzierungsgrad wird durch die Anzahl der Antwortkategorien festgelegt. Krosnick und Presser (2010) kommen in ihrer Studie zu dem Ergebnis, dass fünf bis sieben Antwortkategorien als optimale Lösung gesehen werden können. Mit dieser Anzahl an möglichen Antworten differenzieren Ratingskalen ausreichend und es kommt zu keiner Überforderung durch zu viele Antwortangebote. In den Ratingskalen in dem hier entwickelten Fragebogen wird die fünfte Antwortkategorie mit folgender Begründung weggelassen: Bei einer ungeraden Anzahl möglicher Antworten wird häufig die Mittelkategorie gewählt, nicht nur bei einer neutralen Einstellung, sondern auch aus Gründen des Satisficing und der sozialen Erwünschtheit (vgl. Menold und Bogner 2016, S. 6). Die Testpersonen haben in dem Fragebogen, der vor Beginn des Blogprojekts ausgefüllt wird, die Möglichkeit aus vier vollverbalisierten Antworten einer unipolaren Ratingskala²² zu wählen. Um Unklarheiten beim Verständnis der Items zu vermeiden, werden Stufen der Skala in Textform angegeben und nicht mit numerischen Markern versehen. Die Ratingskala (Items: Stimmt voll und ganz, stimmt eher schon, stimmt eher nicht,

²²Die Skalenpolarität unterscheidet zwischen uni- und bipolaren Ratingskalen. Unipolar bedeutet eine Ratingskala aus einem Kontinuum von einer geringen bis hohen Ausprägung, bipolar bezeichnet eine Ratingskala mit zwei gegensätzlichen Dimensionen (vgl. Menold und Bogner (2016, S. 3)).

stimmt absolut nicht) untersucht in 16 Fragen die persönlichen Einstellungen der Teilnehmenden zu eigenen Interessen und Fähigkeiten im Bereich des Reflektierens, des Bloggens und des Lernens.

Der Fragebogen, der zu Beginn des Blogprojekts digital über den Onlinedienst Google Formulare bearbeitet wird, wird nochmals am Ende des Projekts in gleicher Weise durchgeführt. Die Inhalte des zweiten Fragebogens sind dabei um ein Element ergänzt: Eine zusätzliche Ratingskala mit vier Kategorien und vier Fragen untersucht die Verständlichkeit und unterstützende Funktion der beiden Reflektierhilfen. Die Antworten können in der Auswertung Erkenntnisse über Zusammenhänge zwischen den gezeigten Reflektierfähigkeiten und der eigenen Einstellung gegenüber der Reflektiermethoden aufdecken. Für die Auswertung werden nicht alle erhobenen Daten des Fragebogens verwendet, sondern nur Teile daraus. Der Fragebogen ist im Anhang beigelegt (vgl. Anhang, S. A2). Während der Blog ein qualitatives Erhebungsinstrument darstellt, besteht der Fragebogen aus quantitativen und qualitativen Elementen in Form von offenen und geschlossenen Aufgabenformaten. Für die Datenerhebung werden folglich Methoden aus der traditionellen qualitativen als auch aus der traditionellen quantitativen Forschung verwendet, sodass in dieser Phase eine Methodenkombination besteht.

3.3.3 Sampling

Die Auswahl der zu untersuchenden Personen bzw. Dokumente stellt einen wichtigen Schritt im Forschungsprozess dar (vgl. Flick 2010). Im Unterschied zu quantitativen Forschungsergebnissen haben Ergebnisse der klassisch qualitativen Forschung keinen Anspruch auf statistische Repräsentativität. Jedoch sollen auch dabei über die Analyse von Einzelfällen Hinweise auf allgemeine Erkenntnisse geliefert werden. Generalisierungen sind somit sowohl in quantitative Forschungen als auch qualitativen Forschungen von Bedeutung, jedoch mit unterschiedlichen Eigenschaften und in einem unterschiedlichen Umfang (vgl. ebd.). Eine geeignete Stichprobenziehung zur Ermöglichung von Generalisierungen ist demnach unerlässlich (vgl. Kelle und Kluge 2010, S. 38 ff.). In dieser Studie, die sich am *mixed-methods-design* orientiert und folglich die Kombination quantitativer und qualitativer Forschungsansätze kombiniert, wird eine Stichprobe gewählt, die von ihrer Größe eine qualitative Auswertung noch ermöglicht und gleichzeitig genügend Testpersonen aufweist, um Effektstärken berechnen zu können. Nach Mayring (2015) sollte ein „Stichprobenumfang nach Repräsentativitätsüberlegungen und ökonomischen Erwägungen festgelegt“ (ebd., S. 55) und „nach einem bestimmten Modell gezogen“ (ebd.) werden. Akremi (2019, S. 319 f.) unterscheidet zwischen zwei Grundstrategien von Stichprobenziehungen, der Zufallsstichproben und den nicht-probabilistischen Stichproben. In der quantitativen Sozialforschung werden üblicherweise Zufallsstichproben gewählt, während in der qualitativen Sozialforschung häufig auf nicht-probabilistische Stich-

probenstrategien zurückgegriffen wird (vgl. Häder und Häder 2019). Diese können in zwei Kategorien unterteilt werden: Die willkürliche Auswahl, die auch als *Convenience Sampling* bezeichnet wird (vgl. Flick 1996, S. 88) und die kriteriengeleitete bewusste Auswahlstrategie. Für diese Studie wird sich für letzteres entschieden, auf das im Folgenden genauer eingegangen werden soll: Bei der kriteriengeleiteten bewussten Auswahlstrategie gibt es mehrere Unterformen, die alle das Ziel der theoretischen Generalisierbarkeit verfolgen (vgl. Flick 2010, S. 260). Erwähnt sei hier das theoretische Sampling nach Glaser und Strauß (1967), bei dem die Entwicklung einer Theorie aus Daten im Fokus steht. Die Stichprobenauswahl erfolgt dabei während des Prozesses der Datenerhebung und -auswertung. Sobald weitere Stichproben keine neuen Erkenntnisse mehr bringen, wird das Sampling beendet und es kommt zur theoretischen Sättigung (vgl. ebd.). Die Stichprobe kann bei dieser Strategie folglich nicht am Anfang der Studie definiert werden, sondern entwickelt sich im Laufe des Forschungsverlaufs. Im Falle der hier vorliegenden Studie bietet sich die Strategie nicht an, da die Datenerhebung weitestgehend parallel verlaufen soll, um äußere Einflüsse möglichst konstant halten zu können. Eine weitere Möglichkeit innerhalb der kriteriengeleiteten bewussten Auswahlstrategie ist das selektive Sampling, bei dem ableitend von der Forschungsfrage relevante Kriterien für die Auswahl einer geeigneten Stichprobe entwickelt werden (vgl. Flick 2010; Lamnek und Krell 2016). Die Strategie des selektiven Samplings wird auch für diese Studie gewählt. Für die Untersuchung der Reflektierfähigkeiten von Jugendlichen werden bestimmte Kriterien definiert, die die Auswahl einer geeigneten Stichprobe beeinflussen. Eine Besonderheit des Samplings ist, dass die Schülerinnen und Schüler nicht einzeln ausgewählt werden können, sondern jeweils einzelne Gruppen in Form von Schulklassen über Lehrkräfte für die Stichprobe relevant sind. Folgendes Hauptkriterium ist dabei ausschlaggebend:

- Die Studie soll im Rahmen eines P-Seminars im Rahmen einer BNE stattfinden

Daraus ergeben sich weitere Kriterien, die damit in Verbindung stehen:

- Es sollte sich um Schulklassen der gymnasialen Oberstufe (idealerweise Klasse 11) handeln
- Teilnehmen können nur Schulklassen von bayerischen Schulen
- Die Schülerinnen und Schüler benötigen Zugang zum Internet

Die Kriterien werden anhand theoretischer Überlegungen festgelegt. Zum einen gehen einige Autoren davon aus, dass reflexives Denken erst im Erwachsenenalter vollständig möglich ist (vgl. King und Kitchener 1994, S. 8 f.). Um dem Aspekt gerecht zu werden, wird das Projekt mit älteren, fast erwachsenen Jugendlichen durchgeführt. Weiterhin sollten die Schülerinnen und Schüler Interesse am Thema Nachhaltigkeit haben, da es Hinweise darauf gibt, dass Reflektierprozesse in Verbindung mit Vorwissen und persönlichen Bezug zum Thema stehen und

dies als Grundvoraussetzung erfüllt werden soll (vgl. Brendel 2017, S. 254). Da sich die Bedeutung von Reflektierprozessen im Geographieunterricht auf die Bildungsstandards des Faches und auf den Lernbereich Globales Lernen für Gymnasien bezieht, ist die Studie auf die Schulart Gymnasien beschränkt. Die Beschränkung auf bayerische Schulen ist logistisch zu begründen. Da die Forscherin mehrmals die teilnehmenden Schulklassen besucht, sollten diese leicht erreichbar sein, ebenfalls gibt es dadurch keine Unterscheidung im Lehrplan, der den Unterrichtsverlauf beeinflussen könnte. Der Internetzugang muss sowohl in der Schule als auch Zuhause gewährleistet sein, da die Schülerinnen und Schüler Blogbeiträge online erstellen müssen, die als Grundlage zur Datenauswertung dienen.

Nach der Festlegung der Kriterien werden relevante Schulklassen bzw. deren Lehrkräfte über die Möglichkeit an der Teilnahme dieser Studie informiert. Dafür wird ein Flyer erstellt, der per Post an Gymnasien in der Umgebung verschickt wird. Außerdem wird das Forschungsprojekt über einen E-Mail Verteiler an Geographielehrkräfte verschickt, die bereits Fortbildungen an der Universität Bayreuth besucht haben. Die Aktion startete bereits ein dreiviertel Jahr vor Beginn der Studie, damit die Lehrkräfte noch genügend Zeit haben, bei Interesse ein entsprechendes P-Seminar anzubieten. Insgesamt erhielt die Forscherin fünf positive Rückmeldungen, von denen jedoch abschließend nur drei realisiert werden konnten. Die anderen zwei konnten aufgrund von Nichtzustandekommen des P-Seminars nicht daran teilnehmen. Zwei der teilnehmenden Klassen durchlaufen das Projekt vollständig als Experimentalgruppe. Eine Schulklasse dient als Kontrollgruppe und erhält keine Intervention.

Die Experimentalgruppe setzt sich insgesamt aus 24 Schülerinnen und Schülern zusammen, zehn davon sind weiblich und 14 männlich. Die Experimentalgruppe wird hinsichtlich der Reihenfolge der erhaltenen Reflektierhilfen in zwei Teilgruppen eingeteilt: In der ersten Experimentalgruppe mit Schülerinnen und Schülern, die die Reflektiermethoden in der Reihenfolge Gibbs Reflexivitätszyklus und Hütemethode erhalten, sind fünf weibliche und sieben männliche Testpersonen. Die zweite Experimentalgruppe mit Schülerinnen und Schülern, die die Reflektierhilfen in der Reihenfolge Hütemethode und Gibbs Reflexivitätszyklus erhalten, befinden sich ebenfalls fünf weibliche und sieben männliche Personen. Die Einteilung in diese beiden Gruppen erfolgt geplant mit dem Ziel, möglichst homogene Gruppen zu erzeugen. Orientiert wird sich hierbei am Geschlecht und an den gezeigten Reflektierfähigkeiten der ersten drei Beiträge vor der Einteilung. Die Kontrollgruppe besteht aus 15 Jugendlichen, davon zehn Schülerinnen und fünf Schülern.

3.3.4 Datenerhebung während des Projekts

Die Daten werden während eines Klassenblogprojekts über einen Zeitraum von einem Schuljahr gesammelt. Insgesamt verfasst jede einzelne Schülerin bzw. jeder einzelne Schüler der Experimentalgruppe etwa zehn Blogbeiträge. Während des Projekts wird die Klasse viermal von der Forscherin besucht. Während des ersten Besuchs werden die ersten Daten erhoben, die anschließenden Daten werden außerhalb des Unterrichts erstellt und das letzte Datenmaterial wieder während des Besuchs der Forscherin in der Klasse. Im Folgenden soll der genaue Ablauf des Blogprojekts vorgestellt werden.

Zu Beginn des Projekts und vor dem ersten Schulbesuch wurde von der Forscherin für jede Klasse ein Klassenblog angelegt, auf dem jede Schülerin und jeder Schüler als Autorin bzw. Autor Beiträge unter einem selbst gewählten Pseudonym verfassen kann. Die Schülerinnen und Schüler wählen selbst einen Namen für den Blog und überlegen sich einen Usernamen und ein Passwort für die Anmeldung. Für die Wahl eines sicheren Passworts und eines geeigneten Usernamens, der keine Rückschlüsse auf die tatsächliche Person gibt, erhalten die Jugendlichen ein Informationsblatt. Für die erstmalige Anmeldung im Blog besucht die Forscherin die Klassen vor Ort, um bei technischen Problemen eingreifen zu können. Bevor die Schülerinnen und Schüler mit dem Verfassen der Blogbeiträge beginnen, füllen sie online einen Fragebogen aus. In diesem Fragebogen wird mitunter abgefragt, was sie überhaupt unter dem Begriff *Reflektieren* verstehen und wie ihre persönlichen Einstellungen gegenüber dem Reflektieren sind.

Die erste Blogaufgabe der Projektteilnehmenden ist es, einen Beitrag über Tierschutz zu verfassen. Die genaue Aufgabenformulierung hierzu lautet: „Reflektiere über das Thema Tierschutz und über eine deiner Verhaltensgewohnheiten diesbezüglich“. Das weitgefasste Thema Tierschutz wurde bewusst gewählt, da es ein Thema ist, über das jeder bereits etwas gehört hat und zu dem man je nach individueller Schwerpunktsetzung immer einen Anknüpfungspunkt an die eigene Lebenswelt hat. Außerdem ist es ein Themenfeld, das hinsichtlich des Konzepts des Nachhaltigkeitsdreiecks untersucht werden kann. Als Hilfe für das Verfassen des ersten Beitrags wird mit der Klasse zunächst besprochen, was man in diesem Kontext eigentlich unter Reflektieren versteht. Die Schülerinnen und Schüler erhalten dazu ein Arbeitsblatt mit einer Arbeitsdefinition von Reflektieren (vgl. Anhang, S. A1). Außerdem wird die Aufgabenstellung für künftige Blogaufträge konkretisiert. Die Forscherin betont, dass sich Reflektieren nach dem für das Blogprojekt verwendeten Verständnis stets sowohl auf das Reflektieren über ein bestimmtes Thema als auch auf das Reflektieren über das eigene Verhalten diesbezüglich bezieht. Sie bittet die jeweiligen Lehrkräfte darum die Aufträge immer im gleichen Format mit Hinweis auf beide Reflektierbereiche (Thema selbst + eigene Handlung) zu stellen. Den Tierschutz-Beitrag schreiben die Schülerinnen und Schüler während der Doppelstunde zur Einführung ins Bloggen, bei der die Forscherin anwesend ist. Am Ende der Doppelstunde

veröffentlichen die Schülerinnen und Schüler ihre Beiträge auf ihrem gemeinsamen Klassenblog. Die Privatsphäre der Klassenblogs wurde so eingestellt, dass ihn jeder, der über den Link zum Blog verfügt, besuchen und somit die Beiträge lesen kann. Die Suchmaschinen-Indexierung, d.h. dass der Blog in den Onlinesuchen erscheint, ist deaktiviert. Den Schülerinnen und Schülern ist freigestellt, den Blog auf privat (nur die eigene Klasse inklusive Lehrkraft und Forscherin) oder öffentlich zu stellen. Alle entscheiden sich für die letztere Variante. Jede Blogautorin und jeder Blogautor erhält von der Forscherin eine Projektmappe, in der sich bereits Informationen zum Blogprojekt, eine Anleitung zur Anmeldung im Blog, Hinweise zu Passwörtern und Benutzernamen, und das Blatt mit der Arbeitsdefinition des Begriffs *Reflektieren*. Im Laufe des Blogprojekts wird diese Mappe mit weiteren Materialien für die Reflektierhilfen ergänzt.

Nach der Einführungsstunde ins Bloggen und der Erstellung des ersten Blogbeitrags über Tierschutz, erhalten die Schülerinnen und Schüler von ihrer Lehrkraft die nächsten Blogbeitragsthemen, die idealerweise den im Unterricht behandelten Inhalt aufgreifen. Die Themen sollen in Hinblick auf das Konzept des Nachhaltigkeitsdreiecks untersuchbar sein und einen Bezug zur Lebenswelt der Jugendlichen ermöglichen, wie es Brendel (2017, S. 239) in ihrer Studie empfiehlt. Zur Orientierung hat die Forscherin den Lehrkräften einen Themenpool zur Verfügung gestellt. Die Themen wurden mit Beachtung der UN-Nachhaltigkeitsziele ausgewählt.

Auf jeden veröffentlichten Beitrag – auch dem ersten Beitrag zum Thema Tierschutz – erhalten alle Testpersonen individuelles Feedback von der Forscherin. Dies geschieht in Form von Kommentaren unter den Beiträgen und orientiert sich inhaltlich am für diese Studie festgelegten Begriffsverständnis von Reflektieren. Nach Erlernen der Reflektiermethoden richtet sich das Feedback danach, inwieweit die Methoden umgesetzt wurden und was man gegebenenfalls noch mehr beachten sollte. Das Feedback dient nicht als zusätzliche Intervention, sondern erinnert die Teilnehmenden lediglich an die beiden Teilbereiche des Reflektierens und erinnert an die Aspekte der erhaltenen Reflektierhilfen. Die Entscheidung Feedback zu geben beruht darauf, dass dieses Instrument bei Portfolios ein wichtiger Bestandteil ist (vgl. Kapitel 2.5.1).

Die Lehrkräfte können sich diesem bedienen oder natürlich auch eigene Themen verwenden. Nach zwei Beiträgen, die die Schülerinnen und Schüler zu Hause anfertigen, besucht die Forscherin die jeweilige Klasse erneut. Diesmal wird die Klasse in zwei Gruppen eingeteilt, um je eine Reflektiermethode zu erlernen.

Die Gruppeneinteilung erfolgt wie bereits beschrieben nicht zufällig, sondern nach speziellen Kriterien. Es wird darauf geachtet, dass in den Gruppen eine ähnliche Geschlechteranzahl herrscht und dass in beiden Gruppen sowohl Schülerinnen und Schüler mit höheren als auch geringeren Reflektierfähigkeiten sind. Dieses Kriterium wird anhand der ersten drei Beiträge

(Tierschutz-Beitrag und zwei weitere Beiträge) bestimmt. Weiterhin wird darauf geachtet, dass die Verteilung von Schülerinnen und Schülern, die sowohl über die Inhalte als auch über ihre Handlungen schreiben, und Schülerinnen und Schülern, die überwiegend nur über einen Teilaspekt reflektieren, gleichmäßig ist. Die bewusste Gruppeneinteilung sorgt für die Vergleichbarkeit der Reflektierentwicklungen beider Gruppen, da die Ausgangsbedingungen bei beiden möglichst konstant gehalten werden.

Die erste Gruppe erhält ein Arbeitsblatt mit Gibbs Reflexivitätszyklus als Hilfsmethode, um besser über das eigene Verhalten zu einem vorgegebenen Thema reflektieren zu können. Aufgabe ist es, Beispieltex te mit Hilfe des Reflexivitätszyklus zu analysieren. Die Schülerinnen und Schüler sollen herausfinden, welcher Text ein hohes Reflexivitätsniveau zeigt und dies anhand Gibbs einzelner Phasen begründen. Anschließend ist es die Aufgabe, den eigenen, zuletzt geschriebenen Blogbeitrag zu überarbeiten. Es genügt, wenn die Schülerinnen und Schüler nur den Teil überarbeiten, in dem sie über ihr eigenes Verhalten reflektieren und für den Gibbs Reflexivitätszyklus als Orientierung genutzt werden kann. Der ursprüngliche Reflexivitätszyklus von Gibbs wurde für die Verwendung im Rahmen dieser Studie modifiziert und ins Deutsche übersetzt. Neben einer Aufnahme von Fragen, die direkt Bezug auf das Konzept des Nachhaltigkeitsdreiecks nehmen, wurden die Phasen *evaluation* und *analysis* zu einer zusammengefasst. Die Schülerinnen und Schüler erhalten neben einer Übersicht der fünf zu bearbeitenden Phasen zusätzliche Hilfsfragen, die ihnen beim Durchlaufen der Phasen helfen sollen (vgl. Anhang, S. A4).

Die andere Gruppe wendet gemeinsam mit der Forscherin die Hütemethode nach de Bono zum zuletzt im Blog behandelten Thema an. Der Ablauf ist folgendermaßen gegliedert: Zunächst gibt die Forscherin einen Überblick über die Methode. Dann wird gemeinsam jede Sichtweise der Methode bearbeitet oder anders ausgedrückt, jeder farbige Hut gemeinsam aufgesetzt und besprochen. Zunächst schreibt jeder Teilnehmende für sich die Aspekte auf, die ihm dazu einfallen, anschließend werden diese gemeinsam an der Tafel gesammelt. Nachdem die fünf Sichtweisen (Fakten, Emotionen, positive Aspekte, negative Aspekte, Lösungs ideen) besprochen wurden, wird der blaue Hut, der als metakognitive Kontrollinstanz dient, getragen, um durch das Bewusstsein bzw. Nicht-Bewusstsein bestimmter Perspektiven unterschiedliche Sichtweisen auf ein Thema identifizieren und begründen zu können. Das Konzept des Nachhaltigkeitsdreiecks dient hierbei als Hilfe, um wirtschaftliche, ökologische und soziale Aspekte mit in den Fokus zu nehmen. Als Nachbereitung erhalten die Schülerinnen und Schüler die Aufgabe, ihren zuletzt geschriebenen Blogbeitrag in Hinblick auf de Bonos Multiperspektivität zu untersuchen und sich selbst einen Kommentar zu schreiben, indem sie sich Tipps für das künftige Reflektieren über das Thema selbst geben. Die Hütemethode hat das Ziel, den Schülerinnen und Schülern bewusst zu machen, dass es viele Ansichten zu einem

Thema geben kann, die sich aus unterschiedlichen Vorwissensständen und Erfahrungen ergeben und wodurch ein Thema unterschiedlich bewertet werden kann. Die Hüte dienen dazu, die Sichtweisen zu identifizieren und nachvollziehen zu können. Wichtig ist die Erkenntnis, dass es nicht die eine richtige Perspektive auf ein Thema gibt, sondern dass es viele Meinungen dazu geben kann und es dadurch oft zu Problemen führt. Die Schülerinnen und Schüler sollen aus dieser Erkenntnis heraus ein eigenes Fazit für sich selbst ableiten. Beide Gruppen erhalten jeweils ein Arbeitsblatt mit Anwendungstipps für die erlernte Reflektierhilfe für ihre individuellen Reflektiervorgänge (vgl. Anhang, S. A4 und A7).

Nach dem zweiten Schulbesuch der Forscherin, bei dem jede Gruppe je eine Reflektiermethode erlernt, die künftig beim Verfassen der Blogbeiträge als Unterstützung verwendet werden darf, bearbeiten die Schülerinnen und Schüler wieder drei Blogaufträge, die die Lehrkraft entsprechend zu den behandelten Themen und nach Vorlage der Aufgabenformulierung stellt. Sie haben für jeden Blogbeitrag mehrere Tage Zeit und können ihn flexibel erstellen und frei gestalten. Beim dritten Besuch der Forscherin werden die beiden Gruppen vertauscht, sodass jede Schülerin bzw. jeder Schüler nun die noch nicht erlernte Reflektierhilfen kennenlernt. Zu jeder Reflektiermethode erhalten die Jugendlichen ein Informationsblatt, dass sie in ihre Blogmappe abheften sollen und bei Bedarf zur Hand nehmen können. Anschließend verfassen die Schülerinnen und Schüler wieder ihre individuellen Blogbeiträge zu den aufgegebenen Themen. Der vierte und letzte Schulbesuch der Forscherin dient dazu, gemeinsam mit der ganzen Klasse die beiden Reflektiermethoden zu wiederholen und einen letzten Blogbeitrag zu verfassen. Vor der gemeinsamen Besprechung füllen die Jugendlichen den Online Fragebogen aus, der zu Beginn des Blogprojekts bereits verwendet wurde und durch vier Fragen zu den beiden Reflektierhilfen ergänzt wurde. Die beiden Umfragen dienen dazu, die Ergebnisse vor und nach der Intervention miteinander vergleichen zu können. Das Thema des letzten Beitrags, den die Schülerinnen und Schüler während des letzten Besuchs der Forscherin verfassen sollen, entspricht dem des ersten Blogbeitrags. Sie erhalten die Aufgabe über das Thema Tierschutz und einer ihrer Verhaltensgewohnheiten diesbezüglich zu reflektieren. Dabei können sie sich bei Bedarf an den beiden Reflektierhilfen orientieren. Die Forscherin weist darauf hin, dass sie den Beitrag so schreiben sollen, als hätten sie das Thema zuvor noch nicht im Blog bearbeitet. Die ersten Beiträge zu diesem Thema wurden vorab vom Blog genommen, damit die Schülerinnen und Schüler keinen Zugriff mehr darauf haben. Der letzte Schulbesuch im Rahmen des Projekts findet, wie der erste Schulbesuch auch, im Computerraum der jeweiligen Schule statt, sodass jeder Jugendliche einen eigenen PC für sich zur Verfügung hat und ungestört schreiben kann. Am Ende der Doppelstunde werden die Beiträge wieder auf dem Blog veröffentlicht. Die Kontrollgruppe durchläuft lediglich die Phasen des ersten und letzten Besuchs der Forscherin und dem Verfassen zweier Beiträge zum Thema Tierschutz. Sie

schreiben die Beiträge ebenfalls auf einem Klassenblog, erhalten dazwischen jedoch keine Intervention.

Das Projekt wurde vorab als Vorstudie mit einer 9. Klasse eines bayerischen Gymnasiums im Fach Geographie durchgeführt. Hierbei wurde geprüft, ob die Reflektiermethoden verständlich sind und die Datenerhebung und -auswertung wie geplant möglich ist. Außerdem wurde dadurch sichergestellt, dass das Bloggen für die Jugendlichen keine Probleme darstellt. Die Vorstudie wurde bewusst mit einer Mittelstufenklasse durchgeführt, denn wenn jüngere Teilnehmenden das Bloggen und die Reflektierhilfen verstehen, sollte dies bei den älteren Testpersonen ebenfalls funktionieren. Die Vorstudie verlief wie geplant, sodass am Ablauf nichts mehr verändert wurde und die Voraussetzung für die Hauptstudie, dass der Ablauf ohne technische oder Verständnisprobleme möglich ist, erfüllt ist. Durch den Probelauf zeigten sich kleine Schwachstellen an den beiden Kategoriensystemen, die entsprechend überarbeitet wurden (vgl. Kapitel 3.4.1).

Die beiden Schulklassen der Experimentalgruppe teilen sich wie folgt auf: In Klasse 1 sind zwölf Jugendliche, wovon sechs weiblich und sieben männlich sind. Klasse 2 setzt sich aus zwölf Jugendlichen zusammen, vier davon sind weiblich und acht davon männlich. 22 Testpersonen haben jeweils zehn Beiträge verfasst. Zwei Teilnehmende verfassen nicht alle geforderten Beiträge, können aber mit Berücksichtigung der fehlenden Beiträge für die Auswertung verwendet werden. Eine Testperson ist bei der ersten Teilintervention nicht anwesend, weshalb er für die Teilauswertung der Wirkung der in dieser Stunde erhaltenen Reflektierhilfen ausgeschlossen wird. Beide Klassen bearbeiten insgesamt jeweils neun unterschiedliche Themen, die aus der Perspektive des Konzepts des Nachhaltigkeitsdreiecks untersucht werden können, das Thema Tierschutz wird doppelt behandelt. Folgende Themen, die bis auf das Thema Tierschutz von den Lehrkräften gewählt werden, sind Bestandteil des Blogprojekts (vgl. Tabelle 9):

Tabelle 9 Übersicht der behandelten Themen im Blog

Beitragsnummer	Thema Klasse 1	Thema Klasse 2
1	Tierschutz	Tierschutz
2	Mobilität	Berufe
3	Klimawandel	Verantwortung
erste Teilintervention in Form einer Reflektiermethode		
4	Müll & Entsorgung	Flüchtlinge
5	Lebensmittel	Soziales Engagement
6	Einkaufen	Team
zweite Teilintervention in Form einer Reflektiermethode		
7	Wasser	Reisen
8	Kleidung	Einkaufen
9	Strom	Mobilität
gemeinsame Wiederholung beider Reflektiermethoden		
10	Tierschutz	Tierschutz

Die Hauptdatenquelle, die zur Auswertung herangezogen wird, sind die einzelnen Blogbeiträge der Schülerinnen und Schüler, die im Laufe des Projekts geschrieben werden. Laut Aufgabenformulierung soll immer jeweils zum aufgegebenen Thema selbst als auch über das eigene Verhalten diesbezüglich reflektiert werden, sodass die Beiträge stets nach der gezeigten Reflexivität als auch Reflexion untersucht werden können. Da die Daten bereits in schriftlicher Form vorliegen, müssen sie nicht weiterverarbeitet werden und können so in ihrem Rohformat zur Auswertung genommen werden. Zwischen den Schulbesuchen, bei denen die Jugendlichen die Reflektierhilfen erhalten, sollen die Jugendlichen je drei Blogbeiträge – nach dem ersten Besuch zwei Beiträge – erstellen. Insgesamt liegen idealerweise nach dem Projekt je zehn Beiträge pro Schülerin bzw. Schüler vor, die exakte Anzahl der Beiträge, die für die Auswertung der Interventionsstudie verwendet werden, beträgt 267 (Experimental- und Kontrollgruppe zusammen). Die Blogbeiträge wurden über den Zeitraum eines Schuljahres verfasst und ermöglichen somit eine Untersuchung der individuellen Reflektierentwicklungen der Schülerinnen und Schüler. Im Anhang befindet sich ein Blogeintrag, um einen exemplarischen Einblick in die Textdokumente zu erhalten (vgl. Anhang, S. A9).

3.4 Aufbereitung und Auswertung der Daten

Da das Datenmaterial in dieser Studie bereits in Textform durch das Veröffentlichen von Onlinetexten vorliegt, muss es vor dem Kodierungsprozess nicht transkribiert werden. Die Texte der Teilnehmenden werden lediglich von zusätzlichen, für die Auswertung irrelevanten Material, wie beispielsweise integrierten Fotos oder Videos, bereinigt. Somit kann das erhobene Mate-

rial mit nur minimalen Aufbereitungsschritten für die Kodierung verwendet werden. Die Beitragstexte der Testpersonen werden in die Software MAXQDA eingepflegt und nach Person und Reihenfolge sortiert und mithilfe der skalierenden Strukturierung kodiert. Anschließend erfolgen jeweils Einzelauswertung der Ergebnisse aus der qualitativen Inhaltsanalyse und eine Gesamtzusammenfassung aller Testpersonen. Dabei werden die Schülerinnen und Schüler hinsichtlich ihrer Entwicklungen in Gruppen zusammengefasst. Damit endet der erste Auswertungsschritt. Als Zwischenschritt werden die qualitativen Daten in quantitative Daten umgewandelt und in die Software SPSS eingepflegt. Auch die Daten, die mithilfe des Fragebogens erhoben werden, werden in Excel exportiert und die Aussagen in Textform für eine nachfolgende statistische Auswertung in Form einer Quantifizierung aufbereitet. Den zweiten Auswertungsschritt bildet eine quantitative Auswertung der Daten der Probandinnen und Probanden und zusätzlich der Daten einer Kontrollgruppe, die keine Intervention erfährt. Beide Auswertungsschritte sollen zur Beantwortung der Forschungsfrage beitragen. Im Folgenden wird zunächst der erste Auswertungsschritt in Form der qualitativen Inhaltsanalyse genauer vorgestellt, anschließend wird sich dem quantitativen Auswertungsverfahren gewidmet.

3.4.1 Qualitative Inhaltsanalyse als erster Auswertungsschritt

Zunächst wird allgemein die qualitative Inhaltsanalyse charakterisiert und genauer auf die gewählte Unterform der skalierenden Strukturierung für diese Studie eingegangen. Eine eindeutige Definition zur Inhaltsanalyse gibt es laut Mayring (2015, S. 11 ff.) nicht. Er fasst jedoch Merkmale davon zusammen, die bei den häufigsten Begriffserläuterungen Erwähnung finden. Eine Inhaltsanalyse hat das Ziel, eine „fixierte Kommunikation“ (ebd., S. 13) zu analysieren, dabei systematisch, regelgeleitet und theoriegeleitet vorzugehen, um „Rückschlüsse auf bestimmte Aspekte der Kommunikation zu ziehen“ (ebd.). Er bezeichnet sie folglich auch als „kategoriegeleitete Textanalyse“ (ebd.). Zu den Aufgaben qualitativer Inhaltsanalysen gehören neben der Hypothesenfindung, das als klassisches Element qualitativer Forschung zählt, auch Pilotstudien, Vertiefungen, Einzelfallstudien, Prozessanalysen, Klassifizierungen und die Hypothesenprüfung (vgl. ebd., S. 22 ff.). In dieser Studie stehen die Ziele der Klassifizierung und der Hypothesenprüfung anhand der Untersuchung von Einzelfällen über einen längeren Zeitpunkt hinweg im Vordergrund. Unter Klassifizierung versteht Mayring „die Ordnung eines Datenmaterials nach bestimmten, empirisch und theoretisch sinnvoll erscheinenden Ordnungsgesichtspunkten, um so eine strukturiertere Beschreibung des erhobenen Materials zu ermöglichen“ (ebd., S. 24). Wie auch in dieser Studie können solche Klassifizierungen Basis für quantitative Analysen sein oder sich daraus ergebende Typologien können selbst das Ziel der Analysen sein. Theorien oder Hypothesen können auch innerhalb qualitativer Analysen überprüft werden und bereits ein Einzelfall könnte eine Hypothese widerlegen (ebd., S. 24f.). Ma-

yring merkt an, dass bei der Verwendung der qualitativen Inhaltsanalyse als Auswertungsmethode auf folgende Aspekte besonders geachtet werden muss: Die Bestimmung des Ausgangsmaterials, die Fragestellung und das Ablaufmodell der Analyse mit speziellen qualitativen Techniken (vgl. ebd., S. 54 ff.). Darauf soll im Folgenden mit Bezug auf die hier vorliegende Studie genauer eingegangen werden.

Bei der Bestimmung des Ausgangsmaterials ist es wichtig, eine geeignete Stichprobe für die Datengenerierung zu wählen und diese konkret vorzustellen. Außerdem muss die Entstehungssituation der Analyse genau erläutert und formale Charakteristika des Materials bestimmt werden. Die Stichprobenziehung wurde bereits in Kapitel 3.3.3 ausführlich erklärt. Zusammengefasst bilden in dieser Studie zehn Schülerinnen und 14 Schüler aus zwei Schulklassen die Experimentalgruppe. Insgesamt liegen 237 Blogbeiträge vor, 22 Testpersonen haben jeweils zehn Beiträge verfasst, eine Person hat neun Beiträge und eine weitere acht Beiträge geschrieben. Die Entstehungssituation ist ebenfalls bereits beschrieben worden (vgl. Kapitel 3.3.4) und auch die formalen Charakteristika wurden bereits angesprochen (vgl. Kapitel 3.2). Bei der Fragestellung der Analyse muss sowohl die Richtung der Analyse als auch die theoriegeleitete Differenzierung der Fragestellung berücksichtigt werden. Sprachliches Material bietet unterschiedliche Analyserichtungen: „Man kann den im Text behandelten Gegenstand beschreiben, man kann etwas über den Textverfasser oder die Wirkungen des Textes bei der Zielgruppe herausfinden“ (Mayring 2015, S. 58). In dieser Studie steht die Aussage über den Gegenstand selbst im Vordergrund, konkret die Performanz der Reflektierfähigkeiten. Die Intention der Autorin bzw. des Autors oder die Wirkung auf das Publikum ist dabei nicht von Bedeutung. Die Fragestellung, die theoriegeleitet aus dem aktuellen Forschungsstand und bestehenden Theorien abgeleitet wurde, wurde bereits in Kapitel 3.1 vorgestellt. In dieser Studie soll ausgehend von der Hauptforschungsfrage folgendes untersucht werden: Welche Potenziale bieten ausgewählte Reflektiermethoden für die gezielte Förderung von Reflektierfähigkeiten bestehend aus Reflexions- und Reflexivitätsfähigkeiten von Jugendlichen beim Bloggen im Rahmen einer Bildung für nachhaltige Entwicklung?

Im Folgenden wird zunächst kurz auf unterschiedliche Analyseformen und speziell auf die skalierende Strukturierung eingegangen, bevor anschließend die einzelnen Schritte dieser genauer beschrieben werden. Für Inhaltsanalysen lassen sich drei Grundformen identifizieren: Zusammenfassung, Explikation und Strukturierung. Mayring (2015) beschreibt die Formen wie folgt:

Zusammenfassung: Ziel der Analyse ist es, das Material so zu reduzieren, dass die wesentlichen Inhalte erhalten bleiben, durch Abstraktion einen überschaubaren Corpus zu schaffen, der immer noch Abbild des Grundmaterials ist.

Explikation: Ziel der Analyse ist es, zu einzelnen fraglichen Textteilen (Begriffen, Sätzen, ...) zusätzliches Material heranzutragen, das das Verständnis erweitert, das die Textstelle erläutert, erklärt, ausdeutet.

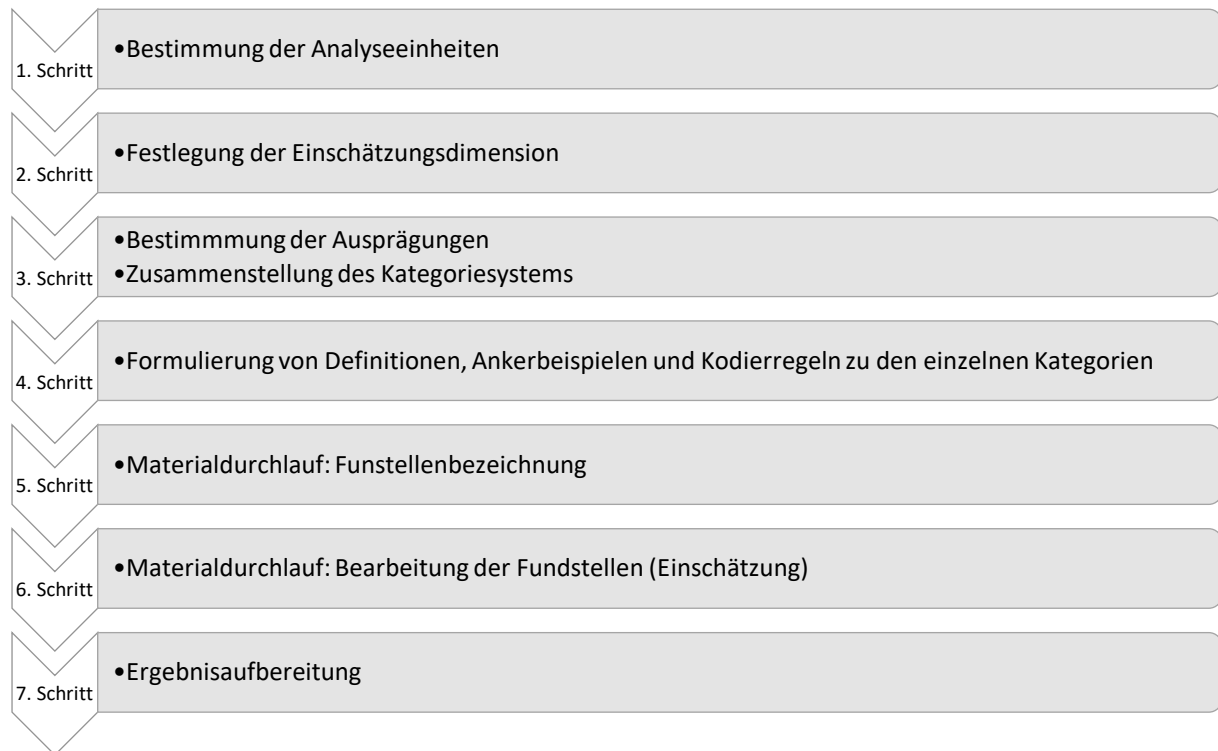
Strukturierung: Ziel der Analyse ist es, bestimmte Aspekte aus dem Material herauszufiltern, unter vorher festgelegten Ordnungskriterien einen Querschnitt durch das Material zu legen oder das Material aufgrund bestimmter Kriterien einzuschätzen. (ebd., S. 67)

Bei der dritten Form, der Strukturierung, wird das Hauptkategoriensystem vorher festgelegt, bei den beiden anderen Formen wird dieses in der Regel erst aus dem Material heraus entwickelt. Alle Grundformen lassen sich in weitere Unterkategorien einteilen, bei der Strukturierung, die für diese Studie relevant ist, wird weiterhin zwischen der formalen, der inhaltlichen, der typisierenden und der skalierenden Strukturierung unterschieden:

Nach formalen Strukturierungsgesichtspunkten kann eine innere Struktur herausgefiltert werden (*formale Strukturierung*). Es kann Material zu bestimmten Inhaltsbereichen extrahiert und zusammengefasst werden (*inhaltliche Strukturierung*). Man kann auf einer Typisierungsdimension nach einzelnen markanten Ausprägungen im Material suchen und diese genauer beschreiben (typisierende Strukturierung). Schließlich kann das Material nach Dimensionen in Skalenform eingeschätzt werden (*skalierende Strukturierung*). (Mayring 2015, S. 68)

Bei allen Formen und vor allem bei der skalierenden Strukturierung, bei der häufig unklar bleibt, wie das Material selbst konkret bearbeitet wird und sich überwiegend auf die Verarbeitung der Ergebnisse fokussiert wird, ist es wichtig, die qualitativen Analyseschritte bei der Erstellung des Kategoriensystems und der Textverarbeitung zu beschreiben (vgl. ebd., S. 106). Um dieser Forderung gerecht zu werden, sollen nun die einzelnen Schritte der skalierenden Strukturierung innerhalb der qualitativen Inhaltsanalyse nach Mayring (2015) für diese Studie genauer vorgestellt werden (vgl. Abbildung 3).

Abbildung 3 Ablaufmodell der qualitativen Inhaltsanalyse in Form der skalierenden Strukturierung nach Mayring (2015)



Bevor die Auswertung durchgeführt werden kann, müssen die Analyseeinheiten festgelegt werden. Mayring (2015, S. 61) unterscheidet dabei zwischen der Kodiereinheit, der Kontexteinheit und der Auswertungseinheit. Die Kodiereinheit definiert dabei den kleinsten, auswertbaren Materialbestandteil, die Kontexteinheit hingegen den größten, auswertbaren Materialbestandteil. Die Auswertungseinheit bestimmt, welche Texte analysiert werden. In dieser Studie entspricht ein Satz der Kodiereinheit und ein sinnzusammenhängender Textabschnitt der Kontexteinheit. Eine Auswertungseinheit bildet jeweils ein Blogartikel. Die Einschätzungsdimensionen sind Variablen mit unterschiedlichen Ausprägungen (vgl. ebd., S. 106). In diesem Fall sind es sowohl die Reflexionsfähigkeiten als auch die Reflexivitätsfähigkeiten der Testpersonen. Beide bestehen aus fünf unterschiedlichen Hauptausprägungen. Die Entwicklung eines Kategoriensystems ist eines der zentralsten Analyseschritte. Bei der skalierenden Strukturierung werden die Kategorien dafür deduktiv entwickelt, d.h. sie werden bereits vor der Auswertung durch Erkenntnisse aus vorherigen Forschungsarbeiten oder der Literatur abgeleitet (vgl. ebd.). Für diese Studie orientiert sich das Kategoriensystem zur Erfassung der Ausprägungen der Reflexionsfähigkeit an Brendels Stufenmodell reflexiven Denkens im Kontext Globalen Lernens (2017) und das Kategoriensystem zur Erfassung der Ausprägungen der Reflexivitätsfähigkeit an den Kompetenzstufen der Reflexivität nach Zimmermann und Welzel (2008). Die Kategorien wurden leicht in Hinblick auf die hier vorliegende Situation angepasst und während

des ersten Durchlaufs im Rahmen einer Vorstudie modifiziert.²³ Im Folgenden sollen die Anpassungen kurz vorgestellt und begründet werden, bevor die beiden Kategoriensysteme inklusive Definitionen, Kodierregeln und Ankerbeispielen präsentiert werden.

Das Stufenmodell reflexiven Denkens im Kontext Globalen Lernens nach Brendel (2017) wurde induktiv aus ihrer Dissertationsstudie heraus entwickelt. In diesem Kontext sollten Schülerinnen und Schüler speziell über Inhalte reflektieren, die sie im Unterricht behandeln. Da die Testpersonen in dieser Studie allgemein über Themen reflektieren sollen, die zwar im Unterricht aufgegriffen werden, um Grundlagen für die Reflexion zu schaffen, jedoch nicht nur Bestandteil der Reflexion sein sollen, wird das Stufenmodell hinsichtlich Elementen, die sich auf konkrete Unterrichtsinhalte beziehen, bereinigt bzw. umformuliert. Nach den ersten Auswertungsversuchen im Rahmen der Vorstudie werden die Stufen 3 und 4 angepasst: Stufe 3 wird in *Begründung/Bewertung* umbenannt und in die Teilbereiche 3a *Begründung* und 3b *Bewertung* eingeteilt. Grund dafür ist, dass die Selbstreflexion eine wichtige Rolle beim Reflektieren einnimmt (vgl. Barth 2018, S. 27 ff.) und sie deshalb mehr gewichtet werden soll als das objektive Reflektieren ohne Bezugnahme zur eigenen Einstellung. Stufe 4 wird in *Analyse* umbenannt, da der Begriff *Beurteilung* ungenau ist, da die Urteilsbildung bereits Bestandteil von Stufe 3 ist. Weiterhin wurden manche Begriffe geändert. Eine Gegenüberstellung des ursprünglichen Stufenmodells von Brendel (2017) und des Kategoriensystems für die Reflexion für diese Studie ist im Anhang zu finden (vgl. Anhang, S. A10).

Das Modell der Kompetenzstufen der Reflexivität nach Zimmermann und Welzel (2008) wird ebenfalls im Rahmen der Vorstudie modifiziert: Die Stufen 1 bis 3 werden in jeweils zwei Unterstufen eingeteilt, wobei sich die erste Unterstufe stets auf das Reflektieren über allgemeines Verhalten bezieht und die zweite Unterstufe auf das Reflektieren über persönliches Verhalten. Dies ist ebenfalls dem Aspekt der Selbstreflexion als wichtiges Element beim Reflektieren geschuldet. Handlungen aus der Ich-Perspektive oder Wir-Perspektive, im Sinne der Familie, gehören zur zweiten Unterstufe und Handlungen aus der allgemeinen Perspektive, wie *man* oder *wir*, im Sinne aller, gehört zur ersten Unterstufe. Weiterhin werden die Kompetenzstufen an den Kontext dieser Studie angepasst und deshalb manche Wörter durch andere Begriffe ersetzt. Das ursprüngliche Modell bezieht sich auf pädagogisches Handeln, dieses wird für den Einsatz in dieser Studie auf professionelles, alltägliches Handeln innerhalb einer BNE adaptiert. Die genauen Umformulierungen und Ergänzungen können der Gegenüberstellung des Modells der Kompetenzstufen der Reflexivität nach Zimmermann und Welzel (2008) und dem Kategoriensystem für die Reflexivität für diese Studie entnommen werden (vgl. Anhang, S. A13).

Die sich aus den bestehenden Modellen abgeleiteten Kategoriensysteme für die Erfassung und Skalierung der Reflexions- und Reflexivitätsfähigkeiten lauten wie folgt:

²³ Die beiden ursprünglichen Modelle werden in Kapitel 2.4.2 ausführlich vorgestellt.

Tabelle 10 Kategoriensystem und Kodierleitfaden für Reflexionsfähigkeiten

Kategorie		Definition	Kodierregeln	Ankerbeispiele
K1: Wiedergabe		<ul style="list-style-type: none"> - Reine Reproduktion von Inhalten oder Gegebenheiten - Faktenwiedergabe - Sachliche Fragen, die anschließend beantwortet werden 	<ul style="list-style-type: none"> - Reine Nennung von Tatsachen oder Fakten ohne Wertung oder Kommentierung/ mit oder ohne vorgelagerter Frage - Wiedergabe von Gehörtem oder Gelesenen 	<p>„121 Liter. Das ist die Menge an Wasser die ein Deutscher an Wasser am Tag verbraucht.“ (P7)</p> <p>„Produkte aus fairem Handel garantieren, dass sie aus einem guten sozialen Klima stammen, das heißt die Arbeiter, die die Produkte herstellen, werden angemessen bezahlt und arbeiten unter menschenwürdigen Konditionen.“ (P11)</p>
K2: Bezug nehmen		<p>Die Schülerin oder der Schüler:</p> <ul style="list-style-type: none"> - äußert eine Beobachtung, Meinung oder Wertung, jedoch ohne Begründung - stellt rhetorische Fragen ohne Antwort - drückt Gefühle aus 	<ul style="list-style-type: none"> - Minimale Wertung von Tatsachen, z.B. durch Wörter wie „natürlich“, „leider“. - Gefühlsäußerung bei Tatsachen - Rhetorische Fragen, auf die keine Antwort gegeben wird, ansonsten K 1 	<p>„Meiner Meinung nach sind Elektroautos eine sehr gute und innovative Erfindung.“ (P3)</p> <p>„Zwar sollten die Eltern, wie es auch bei mir der Fall war, rechtzeitig auf einen zu kommen, und auf das Leben nach der Schule ansprechen, jedoch ist es nicht förderlich dem Schüler bestimmte Berufe oder Studiengänge aufzuzwingen.“ (P24)</p>
K3: Begründung/Bewertung	K3a: Begründung	<p>Die Schülerin oder der Schüler:</p> <ul style="list-style-type: none"> - begründet Prozesse oder Sachverhalte oberflächlich - versucht oberflächlich Wechselbeziehungen darzustellen 	<ul style="list-style-type: none"> - Grund für etwas wird K3a zugeordnet, Faktenaussage/Inhalt dazugehörig K 1 - Signalwörter: weil, aus diesem Grund, deshalb 	<p>„In manchen modernen Ländern herrscht inzwischen auch schon eine Wasserknappheit, weswegen ein Wasserhöchstverbrauch eingeführt wurde.“ (P3)</p> <p>„In 2015 wurde ein Zaun an die Grenze zwischen Ungarn und Serbien gebaut, um zu verhindern, dass Flüchtlinge illegal nach Ungarn kommen.“ (P18)</p>
	K3b: Bewertung	<p>Die Schülerin oder der Schüler:</p> <ul style="list-style-type: none"> - äußert eine Beobachtung, Meinung oder Wertung mit Begründung, Urteil mit simpler Begründung 	<ul style="list-style-type: none"> - Grund für Meinung/Wertung wird K3b zugeordnet, Meinung/Wertung/Beobachtung dazugehörig K2 	<p>„Außerdem kann ich persönlich die Gegner von z.B. Windkraft auch verstehen, die sagen, dass die nicht ständig aus ihrem Fenster auf Windräder schauen wollen, denn schön in der Landschaft sind weder Windräder noch Photovoltaikanlagen.“ (P5)</p>

		<ul style="list-style-type: none"> - stellt persönlichen Zusammenhang zu den Inhalten her, z.B. durch Verbindung mit Vorwissen oder eigenen Erfahrungen - kommt zu Fazit ohne weiter auf vorher genannte Punkte einzugehen 	<ul style="list-style-type: none"> - Wenn eigene Erfahrungen genannt werden, die Meinung/Wertung/Beobachtung beeinflussen, dann K3 - K3 bei einem persönlichen Fazit für einen Aspekt. Das Fazit muss die genannten Punkte nicht nochmal zusammenfassen 	<p>„Klar wäre es auch kein Problem ein paar Euro mehr für Lebensmittel auszugeben, die eigentlich auch viel zu günstig sind. Denn Landwirte auf dem Dorf, die zu ihren Tieren eine viel engere Bindung haben, als die Arbeiter in Großbetrieben, können sich nur schwer gegen die Massenproduktion und die dadurch resultierenden niedrigen Preise, wehren und müssen, deshalb manchmal ihre Betriebe, bei denen der Tierschutz noch groß geschrieben wird verkaufen.“ (P23)</p>
K4: Analyse	<p>Voraussetzung: Konzeptualisierung auf gehobenem Niveau</p> <p>Die Schülerin oder der Schüler:</p> <ul style="list-style-type: none"> - verknüpft Inhalte mit z. B. Theorien oder persönlichen Erfahrungen - führt vertiefte Begründungen für Prozesse, Sachverhalte oder fremde Handlungen an - analysiert ein Problem, sucht nach Antworten und Alternativen, spekuliert oder stellt Hypothesen auf - untersucht den Zusammenhang zwischen Theorie und Praxis - erkennt die Komplexität des Sachverhalts - untersucht Thema aus unterschiedlichen Perspektiven 	<ul style="list-style-type: none"> - K4 erfordert ein gehobenes Abstraktionsniveau - K4 wenn nicht nur eine Begründung für einen Sachverhalt genannt wird, sondern weitere oder eine Begründung tiefer untersucht wird - Problemerkennung und Suche nach Lösungen - Theoretische Annahmen werden mit Praxis in Beziehung gesetzt - Multiperspektivität auf ein Thema wird eingenommen (z.B. durch die drei Bereiche der Nachhaltigkeit) 	<p>„Wirtschaftlich gesehen ist die Methode der schnellen Aufzucht [sic!] durch Mastbetriebe sehr effizient: Man braucht wenig Platz und bekommt mit manipuliertem Futter hohe Erträge auf dem Markt. Tiere werden in der Wirtschaft als Ware oder Gegenstand gesehen und nicht als empfindungsfähige Lebewesen. Diese Thematik ist auch der Ansatzpunkt für viel [sic!] Tierschützer, die sich für bessere Zustände in den Mastbetrieben einsetzen. Persönlich finde ich, dass Mastbetriebe aus der heutigen Konsumwirtschaft und -gesellschaft nicht mehr wegzudenken sind. Also anstatt komplett dagegen zu arbeiten ist die wirkungsvollere Alternative einen Kompromiss zwischen beiden Seiten zu finden. Schlüsselthema hierbei wäre, aufzuhören Tiere ausschließlich als Ressource zu sehen, und den Aspekt des empfindungsfähigen Lebewesens mehr mit einzubringen. Dies könnte sich beispielsweise in neuen Auflagen für Mastbetriebe äußern.“ (P4)</p> <p>„Hart, oder? Das System wie es jetzt ist, ist offensichtlich inhuman und die Hoffnung, dass die junge Generation es revolutionieren könnten habe ich mit meiner Einstellung gerade wahrscheinlich etwas zum Bröckeln gebracht, indem ich genauso egoistisch denke wie alle. Diese Gleichgültigkeit anderen gegenüber ist für den Menschen jedoch bis zu einem gewissen Grad essentiell, damit er ein unbeschwerteres Dasein</p>	

			<p>führen kann, denn wenn ich jeden Tag an das schlimme Schicksal so vieler Menschen in Entwicklungsländern und Industrieländern denken würde, könnte ich mein Leben lang nicht glücklich werden. In einer Welt, die sich nur um Profit dreht und den Menschen als Werkzeug sieht, wir lernen es sogar im schulischen Sozialkundeunterricht im Zusammenhang mit der Funktion der Familie, die die „Gegenwelt zur anonymen, rationalen, funktionalen und unpersönlichen Erwerbswelt [ist]“ (BUR am 22.02.18), kann es keine „Teamarbeit“, in der man sich aufeinander stützen kann und gerne für einander einsteht, geben.“ (P17)</p>
K5: (Re)Konstruktion	<p>Voraussetzung: Konzeptualisierung auf hohem Abstraktionsniveau</p> <p>Die Schülerin oder der Schüler:</p> <ul style="list-style-type: none"> - kommt zu persönlichem Fazit und bezieht systematisch und differenziert Stellung (vorherige Auseinandersetzung mit unterschiedlichen Perspektiven nötig) - formuliert eine persönliche Theorie oder leitet Generalisierungen ab 	<ul style="list-style-type: none"> - hohes Niveau an Abstraktion benötigt für K4 - Fazit + systematische Stellungnahme; beides muss erforderlich sein - Theorieformulierung oder Generalisierung 	<p>„Wir müssen abwägen: Wenn ein hohes menschliches Interesse einem relativ geringen tierischen Interesse gegenübersteht, ist es zu billigen. In Bezug auf die Jagd heißt das, dass die Treibjagd aus Spaß am Jagen verboten sein muss, das Jagen von Wildschweinen zur Regulation des Bestands hingegen akzeptabel ist. Die Aussicht auf die Heilung von Parkinson für Millionen von Menschen rechtfertigt das Leid von tausenden Labormäusen (nach utilitaristischer Auffassung), aber unser Wunsch nach Schönheit kann nicht das Leid von Schimpansen für Kosmetikstudien aufwiegen. Gutes Aussehen ist kein dringendes Lebensinteresse des Menschen und darf daher nicht als solches behandelt werden. Leider ist es oft genau umgekehrt. Ein geringes menschliches Interesse (Spaß an Gänseleberpastete) rechtfertigt ein gewaltiges Interesse des Tieres (Ausbleiben von großen Qualen).“ (P1)</p> <p>„Fazit: Beim Thema Tierschutz gibt es verschiedene Ansichten. Man sollte alle Seiten betrachten und aber auch erkennen, dass sich etwas ändern muss und die Tiere schlecht behandelt werden.“ (P16)</p>

Tabelle 11 Kategoriensystem und Kodierleitfaden für Reflexivitätsfähigkeit

Kategorie		Definition	Kodierregeln	Ankerbeispiele
K1: Sachbezogene Beschreibung	K1a: Objektive, sachbezogene Beschreibung	<ul style="list-style-type: none"> - keine explizite Reflexion - einfache Wiedergabe von Handlungsabläufen - Verallgemeinerung der persönlichen Sicht ohne Begründungen 	<ul style="list-style-type: none"> - Handlungsnennungen allgemein mit man, jeder oder Menschen ohne Begründungen²⁴ - Beschreibung einer Handlung ohne Formulierung der eigenen Perspektive 	<p>„Viele Leute engagieren sich heute ob in Projekten für eine saubere Landschaft, bei Wahlen als Wahlhelfer oder schon die Jugendlichen in der Jugendfeuerwehr.“ (P13)</p> <p>„Wenn wir einkaufen gehen, kaufen wir nicht nur das, was wir brauchen, sondern auch Dinge die interessant aussehen.“ (P6)</p>
	K1b: Subjektive, sachbezogene Beschreibung	<ul style="list-style-type: none"> - keine explizite Reflexion - Beschreibung von Situationen aus der Ich-Perspektive ohne zusätzliche Kommentare, ohne eigene Gedanken, ohne Begründung, ohne Rechtfertigung, ohne Bewertungen 	<ul style="list-style-type: none"> - Beschreibung der eigenen Handlung mit Formulierung der Ich-Perspektive - Handlungsnennungen persönlich mit Ich oder meine Familie/wir ohne Begründungen 	<p>„Ich gehe jede Woche mit meinen Eltern einkaufen und wir gehen meistens in zwei verschiedene Läden.“ (P10)</p> <p>„Nachmittags fahre ich jedoch fast immer mit dem Bus nach Hause.“ (P22)</p>
K2: Handlungsbezogene Begründung	K2a: Handlungsbezogene Begründung für allgemeine Handlung	<ul style="list-style-type: none"> - Beschreibung von Sachverhalten inklusive Begründungen - Bewertungen von allgemeinen Handlungen mit Begründungen - Konkrete Handlungsaufforderung - Nennung von allgemeinen Handlungsvorschlägen/ -alternativen 	<ul style="list-style-type: none"> - zusätzlich zu K1a werden nun Begründungen und ggfl. Auch Wertungen genannt, z.B. durch Formulierungen wie „ich finde es schlecht“, „meiner Meinung nach“ mit Begründungsnennungen wie „weil“ oder „da“ - allgemeine Handlungsvorschläge werden ohne nähere Begründung/Analyse genannt, z.B. durch 	<p>„Immer mehr Leute entscheiden sich dazu Bio-Produkte zu kaufen, da sie wissen, dass die Tiere gut gehalten werden und nicht gequält werden.“ (P16)</p> <p>„Meiner Meinung nach sollte man diese nutzen wenn man die Möglichkeit hat und vor allem in den Städten ist es sehr leicht zu Fuß zu gehen, mit dem Fahrrad zu fahren oder Bus und Bahn zu nutzen.“ (P11)</p>

²⁴ Handlung muss stets selbst durchführbar sein, sonst zählt es nicht zur Reflexivität

			Formulierungen wie „man sollte“, „man könnte“, „es wäre gut, wenn“	
	K2b: Handlungsbezogene Begründung für eigene Handlung	<ul style="list-style-type: none"> - Beschreibung von persönlichen Handlungen mit zusätzlichen Kommentaren, eigenen Gedanken, Begründungen, Rechtfertigungen oder Bewertungen - konkrete Handlungsebene (d.h. keine Verallgemeinerung) - innere Distanzierung vom bloßen Handeln, da Formulierungen der eigenen Perspektive vorhanden - Nennung von persönlichen Handlungsvorschlägen/ -alternativen - künftige Handlungsplanungen ohne Gründe dafür zu nennen 	<ul style="list-style-type: none"> - zusätzlich zu K1b muss eigene Handlung neben Beschreibung auch begründet, kommentiert oder bewertet werden, z.B. durch Formulierungen wie „ich muss zugeben“, „ich muss gestehen“, „weil“, „da“, „aus diesem Grund“ - persönliche Handlungsalternativen werden genannt, z.B. durch Formulierungen wie „ich könnte“, „ich sollte“ - Handlungen für die Zukunft werden genannt ohne weiter darauf einzugehen, z.B. durch Formulierungen wie „in Zukunft werde ich“, „das nächste Mal werde ich“ 	<p>„Ehrlich gesagt bin ich auch jemand der gerne mal eigentlich eher sinnlose Produkte kauft oder Kleidung die man nicht unbedingt dringend benötigt. Prinzipiell finde ich persönlich das auch gar nicht so tragisch, solange es keine Überhand nimmt.“ (P8)</p> <p>„Jedoch bin ich eine Person, die am liebsten alles alleine macht. Aus diesem Grund übernehme ich am liebsten die Rolle des Teamleiters.“ (P18)</p>

K3: Analytische Abstraktion	K3a: analytische Abstraktion für allgemeine Handlung	<p>Sowohl für allgemeine als auch persönliche Handlungen:</p> <ul style="list-style-type: none"> - gemachte Erfahrungen werden unter Berücksichtigung mehrerer Perspektiven/von einer höheren Perspektive aus analysiert - Abwägung und Beurteilung unterschiedlicher Ansätze - ein neues, erweitertes Verständnis wird entwickelt (Fazit aus Überlegungen) - Dialog mit der Situation und mit sich selbst - aus der Analyse heraus werden Konsequenzen im Hinblick auf zukünftiges Verhalten, zukünftige Aktivitäten gezogen 	<p>Sowohl für allgemeine als auch persönliche Handlungen:</p> <ul style="list-style-type: none"> - Handlungen werden aus mind. zwei Perspektiven untersucht (nur aus eigener Perspektive ist K2) - Handlungen werden eingebettet in größeren Kontext, Gründe werden aus allgemeinerer Ebene untersucht - Alternativen werden abgewogen (bei reiner Nennung von Alternativen 2K, bei Begründung dieser und Abwägung mit anderen Handlungen 3K) - Gründe für und gegen Handlungen werden untersucht (bei Nennung jeweils nur eines negativen Aspekts bzw. positiven Aspekts 2K) - Fazit aus Überlegungen wird entwickelt, Zusammenfassung der eigenen Meinung für richtiges allgemeines Handeln - Handlungen für die Zukunft werden genannt inklusive Analyse bzw. Begründungen dafür, ansonsten 2K - Innerer Dialog über Handeln findet statt durch Stellen von Fragen, Ausdrücken von Zweifel 	<p>„Aus meiner jetzigen Situation heraus beurteilt kann ich nur zu gut verstehen, warum sich viele Menschen gar nicht sozial engagieren. Oftmals fehlt einem neben Ausbildung oder Arbeit einfach die Zeit, ehrenamtliche Tätigkeiten irgendwo unterzubringen, sofern man nicht noch ein bisschen Freizeit für sich selbst haben möchte um nach all dem Stress etwas zu entspannen. Jedoch sollte man sich immer dessen bewusst sein, dass es auch keineswegs selbstverständlich ist, dass andere Leute sich für das Gemeinwohl einsetzen. Denn alle freiwilligen Helfer und ehrenamtlichen Arbeiter opfern ihre Zeit, ohne direkt selbst davon zu profitieren. Deshalb finde ich es wirklich bemerkenswert, dass Menschen so sozial sind und beispielsweise in Altenheimen oder Flüchtlingsunterkünften mithelfen.“ (P22)</p>
	K3b: analytische Abstraktion für eigene Handlung			<p>„Bei Produkten aus der Tierzucht mache ich mir meistens keine Gedanken, wie es dem Tier in seinem, durchaus kurzem Leben, ergangen ist. Deshalb muss ich gestehen, dass ich auch ab und zu billiges Fleisch von Aldi oder Lidl kaufe. Zwar bin ich mir dann schon bewusst, dass das Rind oder Schwein kein schönes, gar artgerechtes, Leben hatte, jedoch verleitet mich der günstige Preis dann trotzdem zum Kauf. Einerseits schmecke ich keinen Unterschied zu teurerem Fleisch, andererseits ist es im Gegensatz zu biologischem, artgerechtem gehaltenem Tier, wesentlich schadstoffärmer (vorallem frei von Antibiotika), aber auch teurer.“ (P20)</p>

K4: Kritischer Diskurs	<ul style="list-style-type: none"> - die Erkenntnisse werden im Hinblick auf übergeordnete Werte, Prinzipien und/oder auf relevante Theorien hin bewertet - es erfolgt eine kritische Auseinandersetzung damit (Multiperspektivität Voraussetzung dafür) - Zunahme der Fähigkeit zum Erkennen eigener Anteile und Deutungen in Interaktionen; Bewusstsein, dass frühere Erfahrungen das eigene Handeln beeinflussen; Bewusstsein über eigene Grenzen - Bewusstsein, dass Handeln eingebettet ist in verschiedene Kontexte (z.B. politisch, historisch, gesellschaftlich), die Einfluss nehmen 	<ul style="list-style-type: none"> - Handlung wird analysiert und Erkenntnisse anschließend bewertet mit Kriterien wie Werte oder Prinzipien (ohne Abstrahierung K3) - Gründe für eigenes Handeln wird klar, Metaebene durch eigene Erfahrungen und auch andere Kontexte wird eingenommen (Nicht nur ein Aspekt wird von einer höheren Ebene aus betrachtet, sondern generelles Verhalten) - Grenzen werden erkannt 	<p>„Unser Konsum hat riesige Auswirkungen auf die Umwelt und erzeugt gigantisches Leid in großen Teilen der Welt. Und eigentlich weiß man auch, dass das eigene Handeln moralisch verwerflich ist. Dennoch fühlt man sich selten schlecht oder schuldig beim Kauf von Kleidung. Auch bei mir überwiegt die Freude am Kauf. Bilder von eingestürzten Textilfabriken, von gerupften Angora-Kaninchen und verschmutzten Flüssen machen mich trotzdem betroffen – jedoch nur kurzzeitig. Am nächsten Tag scheint alles wieder vergessen. Ständig ist man mit Leid und Ungerechtigkeit konfrontiert. Es reicht schon Nachrichten zu lesen und man bekommt die Konsequenzen des eigenen Handelns detailliert aufgelistet. Manchmal habe ich den Verdacht, dass mich all dies abstupft. Bisher habe ich noch keinen Ansatz gefunden, richtig damit umzugehen. „Nur ein schlechtes Gewissen ist ein gutes Gewissen“ war ein Wahlspruch der 68-Bewegung. Vielleicht ist da etwas dran.“ (P1)</p>
K5: Professionalisierung	<ul style="list-style-type: none"> - Erkennen des eigenen Urteils-Handlungs-Hiatus UND Durchdenken, Planen von zielorientierten Verbesserungsstrategien - konkrete Verbesserungsmaßnahmen (am eigenen Tun) werden initiiert - festgefahrene Rollenverständnisse werden thematisiert und konstruktiv in Frage gestellt, andere Handlungsmöglichkeiten diesbezüglich ausgelotet und ausprobiert 	<p>-</p>	<p>- <i>nicht vorhanden</i> -</p>

Die Kategoriensysteme, die zunächst auf Grundlage der beiden Stufenmodelle entwickelt und durch die Anwendung in einer Vorstudie modifiziert wurden, dienen als Kodierleitfaden für den Auswertungsprozess. Neben den Kategorien und den entsprechenden Definitionen bestehen sie aus Kodierregeln und Ankerbeispielen, die während des Auswertens erweitert und angepasst werden. Für die Zuordnung von Textstellen in geeignete Kategorien hat sich das Verfahren mit den drei Merkmalen der einzelnen Kategorien bewährt (vgl. Ulich et al. 1985). Bei beiden Kategoriensystemen muss jeweils nur eine Kodierregel erfüllt sein, damit die entsprechende Kategorie erreicht wird. Die Kriterien für niedrigere Kategorien müssen dabei nicht zwingend als Voraussetzung erfüllt sein. Der Materialdurchlauf gliedert sich in zwei Schritte: Zunächst werden die Fundstellen innerhalb des Textes mit Notierung der entsprechenden Kategoriennummer markiert, anschließend werden die Fundstellen je nach Ziel der Strukturierung bearbeitet und aus dem Text extrahiert (vgl. Mayring 2015, 97 ff.). Im Folgenden wird beispielhaft die Auswertung eines Blogbeitrages in Form der skalierenden Strukturierung mit den Einschätzungsdimensionen Reflexionsfähigkeit und Reflexivitätsfähigkeit dargestellt:

Abbildung 4 Kodierter Text mit Anwendung beider Kategoriensysteme zum Thema Einkaufen von P3



Beim Auswerten wird folgendermaßen vorgegangen: Das Material wird von zwei Personen zunächst mit dem Kategoriensystem für die Reflexivitätsfähigkeit, anschließend mit dem für die Reflexionsfähigkeit kodiert, wobei jede Person eigenständig arbeitet. Die roten Textstellen stellen Einordnungen in das Kategoriensystem zur Bestimmung der Reflexivitätsfähigkeit dar, die Grünen in das Kategoriensystem zur Bestimmung der Reflexionsfähigkeit. Textpassagen, die sowohl Reflexions- als auch Reflexivitätsanteile enthalten werden doppelt kodiert. Bei regelmäßigen Treffen der Kodierer werden die Ergebnisse miteinander verglichen und Stellen mit keiner Übereinstimmung genauer untersucht und sich aus der Diskussion heraus auf eine Kategorienzuordnung geeinigt. Die Zweitkodierung wird nochmal ausführlicher in Kapitel 3.4.2

behandelt, da sie ein wichtiges Element der Sicherung der Gütekriterien in dieser Studie darstellt.

Der letzte Schritt beim Analyseverfahren entspricht der Ergebnisaufbereitung. „Bei der Ergebnisaufbereitung [...] werden diese Einschätzungen zusammengefasst und schließlich nach Häufigkeiten, Kontingenzen oder Konfigurationen quantitativ analysiert.“ (Mayring 2015, S. 106). In dieser Studie wird zunächst pro Person eine Einzelauswertung durchgeführt, anschließend werden die Ergebnisse zusammengefasst. Die Einzelauswertung gliedert sich dabei in folgende Untersuchungsbereiche:

1. Allgemeines zur Person
 - a. Personenbezogene Merkmale
 - b. Begriffsverständnis von Reflektieren
 - c. Bewertung der Reflektierhilfen
2. Vergleich der Beiträge zum Thema Tierschutz
 - a. Score
 - b. Analyse der Reflektierstufen
3. Vergleich der Beiträge vor, zwischen und nach den Reflektiermethoden
 - a. Score
 - b. Analyse der Reflektierstufen

Im Folgenden werden die einzelnen Gliederungspunkte der Einzelauswertung genauer beschrieben. Der erste Punkt dient dazu einen Überblick über die jeweilige Testperson zu erhalten. Er enthält Informationen über das Geschlecht, die Klassenzugehörigkeit, die Anzahl an geschriebenen Beiträgen während des Projekts und die Reihenfolgen, in der die Person die Reflektierhilfen erhält. Weiterhin werden die Begriffsverständnisse von Reflektieren der Schülerin bzw. des Schülers am Anfang und am Ende des Projekts vorgestellt und miteinander verglichen. Zuletzt wird dargestellt, wie hilfreich und verständlich die Testperson die Reflektiermethoden selbst bewertet. Die Informationen für die Definitionen und die Selbsteinschätzung können dem Online-Fragebogen, der zu Beginn und am Ende des Projekts von den Teilnehmenden ausgefüllt wurde, entnommen werden.

Der zweite Punkt soll untersuchen, inwieweit sich die Reflektierfähigkeiten der entsprechenden Testpersonen durch die beiden Reflektierhilfen insgesamt verändern. Die Schülerinnen und Schüler verfassen jeweils zwei Beiträge zum Thema Tierschutz. Der erste Beitrag wird nach einer kurzen Einführung in das Bloggen und nach Klärung des Begriffs *Reflektieren* während des Unterrichts von jedem einzeln verfasst (vgl. Kapitel 3.3.4). Der zweite Beitrag zum gleichen Thema wird am Ende des Blogprojekts und folglich nach dem Erlernen beider Reflektierhilfen ebenfalls während des Unterrichts geschrieben, nachdem gemeinsam mit den Schülerinnen und Schülern die beiden Reflektiermethoden wiederholt wurden. Die beiden Beiträge können

gut miteinander verglichen werden, da sie über das gleiche Thema handeln und die Aufgabenstellung identisch ist und lautet: Reflektiere über das Thema Tierschutz und über eine deiner Verhaltensgewohnheiten diesbezüglich. Beim Verfassen des zweiten Beitrags zu diesem Thema können die Testpersonen ihren ersten Beitrag nicht mehr einsehen. Sie werden dazu aufgefordert, so zu schreiben, als würden sie erstmalig über das Thema im Blog reflektieren. Durch den Vergleich der beiden Beiträge zum Thema Tierschutz kann untersucht werden, welche Potenziale die beiden Reflektierhilfen insgesamt auf die Reflexions- und Reflexivitätsfähigkeit haben und wie sich die Qualität des Reflektierens dadurch verändert. Beide Beiträge werden in der Schule geschrieben, wodurch gleiche Bedingungen garantiert werden und die Zeit, die für die Beiträge zur Verfügung gestellt wird, abgestimmt werden kann.

Neben den beiden Beiträgen zum Thema Tierschutz, die im Unterricht verfasst werden, schreiben die Schülerinnen und Schüler weitere Beiträge, die sie jeweils während einer gewissen Zeit (i.d.R. zwei Wochen) zuhause fertigstellen. Diese werden in den Einzelauswertungen im dritten Punkt behandelt. Die Beiträge handeln von verschiedenen Themen, die die Lehrkraft festlegt und die zuvor im Unterricht behandelt wird. Vor Erlernen der ersten Reflektiermethode verfassen alle Schülerinnen und Schüler zwei Beiträge. Die Mittelwerte der qualitativen Daten der einzelnen Beiträge bildet den ersten Textblock, der für die Untersuchung herangezogen wird. Nach Erhalt der ersten Reflektierhilfe werden drei weitere Beiträge zuhause geschrieben. Der Mittelwert der qualitativen Daten dieser drei Beiträge formt den zweiten Textblock und der Mittelwert der qualitativen Daten der weiteren drei Beiträge, die zwischen Erhalt der zweiten Reflektierhilfe und dem Verfassen des letzten Beitrags zum Thema Tierschutz im Unterricht zuhause geschrieben werden, bilden den dritten Textblock. Idealerweise besteht der erste Textblock folglich aus den Mittelwerten der Ergebnisse von zwei Beiträgen und der zweite und dritte Textblock aus den Mittelwerten der Ergebnisse von je drei Beiträgen. Die Untersuchung der Potenziale der einzelnen Reflektiermethoden wird durch Vergleich der Textblöcke vor und nach der entsprechenden Reflektierhilfe durchgeführt. Die Zusammenlegung der einzelnen Beiträge zu einem Mittelwert ist dadurch begründet, dass die Beiträge unterschiedliche Themen behandeln und hierbei unterschiedlich viel Interesse und Motivation gezeigt wird und dadurch die gezeigte Reflektierfähigkeit beeinflusst werden könnte.²⁵ Ebenfalls werden die Beiträge zuhause zu unterschiedlichen Zeitpunkten verfasst und es bleibt ungewiss, wieviel Zeit die Probandinnen und Probanden in die einzelnen Schreibphasen tatsächlich investieren konnten und wollten. Da keine gleichen Voraussetzungen für das Entstehen der einzelnen

²⁵ Nach aktuellem Wissenstand der Autorin gibt es aktuell noch keine Erkenntnisse über Zusammenhänge von Motivation und Interesse auf das Reflektieren. Da Motivation und Interesse jedoch deutliche Einflüsse auf das Lernen haben (Krapp (2010)), soll die Bedeutung dieser beiden Faktoren für diese Studie ebenfalls berücksichtigt werden.

Beiträge gewährleistet werden können (im Gegensatz zu den Beiträgen zum Thema Tierschutz), ist der Vergleich der einzelnen Beiträge miteinander nicht zielführend. Es ist viel mehr von Interesse herauszufinden, wie die Reflektierfähigkeiten sich im Durchschnitt über einen längeren Zeitraum hinweg entwickeln und ob durch die Reflektierhilfen neue Reflektierperformanzen gezeigt werden.

Für die Analyse der Reflektierstufen werden die Häufigkeiten der Kategorien aus der qualitativen Inhaltsanalyse verwendet. Dabei werden die prozentualen Verteilungen auf die einzelnen Kategorien vorgestellt und zusätzlich die am häufigsten und am höchsten erreichten Kategorien je Kategoriensystem benannt und mit den Ergebnissen der anderen Beiträge verglichen. Die Darstellung der Ergebnisse mit Nennung der Häufigkeiten wurde beispielsweise bereits in den Studien von Brendel (2017) und Eysel (2006) als Ergebnisdarstellung durchgeführt. Damit die Beiträge besser miteinander verglichen werden können, werden zusätzlich dazu aus den qualitativen Daten Scores für die Gesamtreflektierfähigkeit, die Reflexionsfähigkeit und die Reflexivitätsfähigkeit berechnet. Dieses Verfahren beinhaltet den Schritt der Quantifizierung qualitativer Daten, bei dem in der *mixed-methods* Forschung „qualitative Daten mit quantifizierenden Techniken in numerische Daten transformiert werden“ (Vogl 2017, S. 291). Somit können Unterschiede der Reflektierfähigkeiten zwischen Beiträgen besser veranschaulicht und kommuniziert werden, da Quantifizierungen der Vereinfachung der Darstellung komplexer qualitativer Daten dienen (vgl. Sandelowski et al. 2009, S. 210). Dieser zusätzliche Schritt soll helfen, die Entwicklung der Reflektierfähigkeiten multidimensionaler zu beleuchten.

Im Folgenden soll die Berechnung der Scores beschrieben werden: Der Gesamtscore im Reflektieren pro Beitrag besteht aus zwei Teilscores, dem Reflexionsscore und dem Reflexivitätsscore eines Beitrags. Die Scores berechnen sich dabei wie folgt: Jeder Reflektierstufe wird ein fester Wert zugewiesen. Dieser Wert wird mit dem prozentualen Anteil der entsprechenden Kategorie am Text multipliziert und mit den weiteren Kategorienscores zu einem Score berechnet. Die Werte der Reflektierkategorien nehmen von Kategorie 1 bis Kategorie 5 um jeweils den Faktor 10 pro Kategorie zu. Unterkategorien, wie beispielsweise K1b oder K2b werden mit einem Zwischenwert versehen und sind um den Faktor 5 höher als die Kategorien a, die als volle Kategorie bzw. Reflektierstufe gezählt werden (vgl. Tabelle 12 und Tabelle 13). Grund dafür ist der, dass die Kategorien b sich auf das eigene Ich beziehen und dies ein höheres Niveau im Reflektieren darstellt, als allgemeines nicht auf die eigene Person bezogenes Reflektieren, da das Selbst im Fokus steht (vgl. Barth 2018, 27 ff.).

Tabelle 12 Grundlage Scoreberechnung für Reflexionsfähigkeit

Kategorie	Wert für Scoreberechnung
K1	10
K2	20
K3a	30
K3b	35
K4	40
K5	50

Tabelle 13 Grundlage Scoreberechnung für Reflexivitätsfähigkeit

Kategorie	Wert für Scoreberechnung
K1a	10
K1b	15
K2a	20
K2b	25
K3a	30
K3b	35
K4	40
K5	50

Folgendes Beispiel soll die einzelnen Schritte für die Berechnung der drei Scores veranschaulichen: Eine Testperson weist in einem Beitrag in Hinblick auf die Reflexionsfähigkeit die Kategorien K1 mit 22%, K2 mit 29% und K3b mit 10% Anteil am Gesamttext auf. Der Reflexionscore für den gesamten Text ergibt insgesamt 11,5 und setzt sich aus folgender Formel zusammen:

Reflexionsscore =

Vorkommen der Kategorie 1 in Prozent x Wert 1

+ Vorkommen der Kategorie 2 in Prozent x Wert 2

+ Vorkommen der Kategorie 3a in Prozent x Wert 3a

+ Vorkommen der Kategorie 3b in Prozent x Wert 3b

+ Vorkommen der Kategorie 4 in Prozent x Wert 4

+ Vorkommen der Kategorie 5 in Prozent x Wert 5

Eingesetzt mit den Werten der Beispielperson sieht die Formel wie folgt aus:

$$22\% \times 10 + 29\% \times 20 + 0\% \times 30 + 10\% \times 35 + 0\% \times 40 + 0\% \times 50 = 11,5$$

Nach diesem Prinzip wird auch der Score für die Reflexivität berechnet. Der Reflektierscore insgesamt ergibt sich durch die Addition der beiden Teilscores. Durch die Scores kann die Entwicklung der Reflektierfähigkeiten über die Laufzeit des Projekts hinweg untersucht werden. Zum einen werden die Scores der beiden Beiträge zum Thema Tierschutz miteinander verglichen, zum anderen werden die Mittelwerte der Scores der einzelnen Beiträge innerhalb der Textblöcke verwendet, um Aussagen über die Potenziale der einzelnen Reflektierhilfen treffen zu können. Die Scoreberechnung nach diesem Verfahren wird dadurch gerechtfertigt, dass dem Kategoriensystemen Stufenmodelle fürs Reflektieren zu Grunde liegen, deren unteren Stufen geringere Reflektierfähigkeiten aufweisen als die höheren Stufen. Folglich werden den niedrigeren Kategorien geringere Werte für die Scoreberechnung zugeschrieben, als den

höheren Kategorien. Die Wert-Abstände zwischen den vollen Stufen bleiben gleich, da davon ausgegangen wird, dass sich die Qualitäten innerhalb der einzelnen Modelle gleichmäßig voneinander unterscheiden. Ähnliche Verfahren zur Auswertung von Reflektierfähigkeiten mit der Berechnung von Scores wurden bereits in anderen Studien, wie beispielsweise bei Chen et al. (2009) durchgeführt. Die Ergebnisse der qualitativen Inhaltsanalyse werden durch Einzelauswertungen (vgl. Kapitel 3.5.1) und einer Zusammenfassung der Ergebnisse durch die Gruppierung der Testpersonen hinsichtlich ihrer Entwicklungen der Reflektierfähigkeiten (vgl. Kapitel 0) dargestellt.

3.4.2 Gütekriterien der Inhaltsanalyse

Die Übertragbarkeit der klassischen Gütekriterien Reliabilität, Validität und Objektivität auf inhaltsanalytische Forschungen wird häufig kritisiert (vgl. Lisch und Kriz 1978, S. 90). Deshalb gibt es Vorschläge für eigene inhaltsanalytische Gütekriterien, wie beispielsweise die acht Konzepte nach Krippendorff (1980). Er schlägt folgende Kriterien für die Güte von Inhaltsanalysen vor: Semantische Gültigkeit, Stichprobengültigkeit, korrelative Gültigkeit, Vorhersagegültigkeit, Konstruktgültigkeit, Stabilität, Reproduzierbarkeit und Exaktheit (vgl. ebd., S. 158). In dieser Studie wird der Fokus auf das Konzept der Reproduzierbarkeit gelegt, das den Grad bezeichnet, „in dem die Analyse unter anderen Umständen, anderen Analytikern zu denselben Ergebnissen führt“ (Mayring 2015, S. 127). Sie wird durch die Explizitheit und Exaktheit der Beschreibung der Vorgehensweise beeinflusst und kann durch die Intercoderreliabilität bestimmt werden. Krippendorff (1980, S. 133 ff.) schlägt dafür einen Koeffizienten vor, um die Übereinstimmung der Kategorienzuordnung mehrerer Kodierer messen zu können. Da in dieser Studie das gesamte Material von zwei Kodierern analysiert wird und Unstimmigkeiten während einer inhaltlichen Diskussion behandelt werden, ist die Berechnung eines Koeffizienten in diesem Fall nicht nötig. Die Einigung ist hier der beste Weg zur Reliabilität, der Koeffizient wäre nur nötig, falls es keinen gemeinsamen Austausch über die Auswertung hinweg gäbe. Obwohl die Zweitkodierung am gesamten Material einen erheblichen Zeitaufwand bedeutet, wurde sich in dieser Studie bewusst dafür entschieden. Grund ist u.a., dass an qualitativen Studien oft kritisiert wird, die Ergebnisse seien zu einem hohen Grad durch subjektive Interpretationen beeinflusst und folglich nicht nachvollziehbar. Mit einer vollständigen Intercoderreliabilität soll dieser Kritik entgegengewirkt und dem Gütekriterium der Reliabilität gerecht werden. Die Sicherung einer inhaltsanalytischen Validität findet in dieser Studie durch das Konzept der Konstruktvalidität statt. Es besagt, dass ein Erhebungsverfahren zur Messung eines Konstrukts in ein theoretisches Setting eingebettet sein und mit etablierten Modellen übereinstimmen sollte. Das zu messende Konstrukt muss vorab genau definiert und aus theoretischen Erkenntnissen hergeleitet werden (vgl. Mayring 2015, S. 127). In diesem Fall kann die Mes-

sung der Reflektierfähigkeiten als konstruktvalide angesehen werden, da die Konstrukte Reflexion und Reflexivität zunächst inhaltlich definiert und voneinander abgegrenzt werden und sich ihre messbaren Ausprägungen aus etablierten Modellen ableiten. Die Untersuchung der Konstrukte wird hierbei nicht willkürlich vorgenommen, sondern unterliegt theoretischen Fundierungen, die in Kapitel 2.4.2 bereits ausführlich vorgestellt wurden.

Während das Konzept der Reproduzierbarkeit nach Krippendorff (1980, S. 158) ein Kriterium für die Reliabilität einer Studie ist, kann das Konzept Konstruktgültigkeit als Kriterium für die Validität im engeren Sinne gesehen werden. Durch die Fokussierung auf die beiden Konzepte – der Interoderreliabilität wird dabei am meisten Gewichtung gegeben – wird in dieser Studie die Sicherung der Güte gewährleistet.

3.4.3 Quantitative Auswertung als zweiter Auswertungsschritt

In diesem Abschnitt wird der zweite Teil der Auswertung genauer vorgestellt. Im Rahmen einer quantitativen Auswertung werden vordefinierte Hypothesen durch Varianzanalysen geprüft, um Antworten auf die (Teil-)Forschungsfrage(n) geben zu können. Zunächst wird das gewählte Auswertungsverfahren ausführlicher vorgestellt und begründet, bevor ein Überblick über die Hypothesen gegeben wird.

3.4.3.1 Begründung des ausgewählten Auswertungsverfahrens

Für die weitere, statistische Auswertung der Daten, die mit Hilfe der qualitativen Inhaltsanalyse nach Mayring (2015) gewonnen werden, finden Varianzanalysen Anwendung. Durch Varianzanalysen werden Unterschiede abhängiger Variablen zwischen verschiedenen Gruppen auf Signifikanz untersucht. Unabhängige Variablen werden dabei Faktoren genannt und die einzelnen Ausprägungen davon als Stufen bezeichnet (vgl. Janczyk und Pfister 2013, S. 137). Die manuelle Berechnung innerhalb einer Varianzanalyse besteht nach Hubert et al. (2014) aus sechs Schritten:

1. Berechnung der quadrierten Gesamtsummenabweichung
2. Berechnung der quadrierten Abweichungen zwischen den Gruppen
3. Berechnung der quadrierten Abweichungen innerhalb der Gruppen
4. Berechnung der relevanten Freiheitsgrade
5. Berechnung der mittleren quadrierten Abweichungen
6. Berechnung und Vergleich der F-Wert (ebd., S. 44)

Die einzelnen Berechnungsschritte werden mit Hilfe statistischer Software, wie beispielsweise dem Programm SPSS vom Hersteller IBM, das auch in dieser Studie Anwendung findet, automatisch durchgeführt. Die Ergebnisse werden in tabellarischer Form angezeigt und das Signifikanzniveau und weitere Daten können so entsprechend ausgelesen werden.

Der Ursprung der Varianzanalyse liegt bei Sir Ronald Aylmer Fisher im 20. Jahrhundert: „Er versteht dieses Verfahren im Sinne einer Abtrennung solcher Varianzen, die auf bestimmte

Ursachen zurückführbar sind, von den übrigen Varianzen, deren Ursachen nicht klar zu bestimmen sind“ (Rasch et al. 2014, S. 5). Elementar für die Varianzanalyse ist folglich das Prinzip der Streuungszerlegung. Die Gesamtstreuung einer Variable kann in eine erklärte und eine nicht-erklärte Streuung zerlegt werden. Die erklärte Streuung bezieht sich auf die Streuung zwischen Gruppen, die nicht-erklärte Streuung auf die Streuung innerhalb von Gruppen. Indem man beide Streuungsgrößen zueinander in Relation setzt, erhält man die sogenannte Testgröße F der Varianzanalyse, die passend zum vergebenen Buchstaben F verteilt ist. Der theoretische F -Wert, der abhängig vom Signifikanzniveau und den dazugehörigen Freiheitsgraden ist, findet sich in den Tabellen der F -Verteilung. Fällt die Testgröße F größer aus als der theoretische F -Wert, so wird die Nullhypothese²⁶ abgelehnt (vgl. Bamberg et al. 2017).

Da die Varianzanalyse sehr allgemein gehalten ist, da sie nur prüft, ob überhaupt signifikante Zusammenhänge zwischen unabhängigen und abhängigen Variablen bestehen und keine Aussagen zu marginalen Wirkungen trifft, wird sie häufig als Basisinstrument für die Analyse von Studien verwendet. Darauf aufbauend folgen gewöhnlich post-hoc Tests und häufig auch Effektgrößenberechnungen, um gezieltere Aussagen über Unterschiede abhängiger Variablen ableiten zu können (vgl. Bamberg et al. 2017). In manchen Studien finden auch Kontrastanalysen Anwendung, wenn bereits vorher Hypothesen zu den Unterschieden zwischen den Gruppen bestehen. Dies ist jedoch eher selten der Fall, weshalb das häufigste post-hoc Verfahren spezielle post-hoc Tests sind (vgl. Field 2013; Bühl 2012). Auch in dieser Studie werden post-hoc Tests im Anschluss durchgeführt, da zwar vermutet wird, dass die Intervention Effekte auf die abhängigen Variablen zeigt, aber nicht in welche Richtung und bei welchen abhängigen Variablen diese genau auftreten. Es existiert eine Vielzahl an möglichen post-hoc Tests für Varianzanalysen, wie beispielsweise den Tukey Test, den Q nach R-E-G-W Test, den LSD Test, den Gabriel Test oder den Games-Howell Test. Die Tests können hinsichtlich der Kriterien Strenge der Kontrolle des Fehlers 1. Art, Strenge der Kontrolle des Fehlers 2. Art²⁷ und Robustheit des Tests in Hinblick auf Verletzungen der Voraussetzungen beurteilt werden (vgl. Keller 2015). Field (2013, S. 472) empfiehlt je nach Gegebenheit folgende post-hoc Tests: Bei gleichen Varianzen und gleicher Stichprobenzahl den Q nach R-E-G-W Test, bei gleichen Varianzen und minimal unterschiedlicher Fallzahl den Gabriel Test, bei gleichen Varianzen und stark unterschiedlicher Stichprobenzahl den GT2 nach Hochberg Test und bei ungleichen Varianzen generell den Games-Howell Test. Da post-hoc Tests keine Ergebnisse über die Stärke des Einflusses einer oder mehrerer Faktoren auf abhängige Variablen liefern, werden häufig außerdem Effektstärken als zusätzliche Komponente ausgewertet (vgl. Hubert

²⁶ Die Nullhypothese nimmt keine signifikanten Unterschiede zwischen zwei oder mehreren Gruppen bzw. zwei Merkmalen an. Alternativhypothesen definieren das logische Gegenteil solcher Annahmen (Reinders et al. (2015, S. 136).

²⁷ Fehler 1. Art bezeichnet „[d]ie unberechtigte Ablehnung der Nullhypothese“ (Sachs (1997, S. 184). Fehler 2. Art bezeichnet [d]as unberechtigte Beibehalten der Nullhypothese“ (ebd.).

et al. 2014, S. 86). Für Effektstärken liegen unterschiedliche Maße vor, wie beispielsweise das partielle Eta-Quadrat oder Cohens d . Cohen definiert drei Schwellenwerte für Effektstärken, die allgemein anerkannt sind (vgl. Cohen 1988, S. 280 ff.). Für Cohens d , das für diese Studie verwendet wird, spricht man bei Werten unter .020 von keinem Effekt, bei Werten zwischen .200 und .400 von einem kleinen Effekt, bei Werten zwischen .500 und .700 von einem mittleren Effekt und bei Werten ab .800 von einem großen Effekt (vgl. ebd.).

Varianzanalysen sind wie auch t-Tests Teil der Dependenzanalysen. Durch ein solches Verfahren können Abhängigkeiten metrisch skalierten Variablen von nominalen Variablen analysiert werden (vgl. Hammann und Erichson 2000, S. 318). Die Entscheidung, die Daten dieser Studie mit Hilfe von Varianzanalysen anstelle von t-Tests auszuwerten, ist dadurch begründet, dass sie eine Erweiterung von t-Tests darstellen und bei t-Tests höchstens Variablenmittelwerte zweier Gruppen bzw. zwischen zwei Messzeitpunkte miteinander verglichen werden können. In dieser Studie werden zum Teil drei Gruppenergebnisse gegenübergestellt und weiterhin drei Messzeitpunkte berücksichtigt, wofür Varianzanalysen benötigt werden (vgl. Janczyk und Pfister 2013, S. 137). Außerdem wirken Varianzanalysen der α -Fehler-Kumulierung entgegen, diese Tatsache ist bei dem wiederholten Anwenden von t-Tests nicht gegeben (vgl. Rasch et al. 2014, S. 66).

Dadurch, dass die Anforderungen von Varianzanalysen gering ausfallen, finden sie häufig Anwendung bei Experimenten, um Hypothesen zu überprüfen. Man unterscheidet zwischen univariaten n-faktoriellen Varianzanalysen und multivariaten n-faktoriellen Varianzanalysen. Ersteres untersucht die Wirkung einer oder mehrerer unabhängiger Variablen auf eine abhängige Variable und letzteres fokussiert sich auf die Wirkung einer oder mehrerer unabhängiger Variablen auf mehrere abhängige Variablen (vgl. Backhaus et al. 2000, S. 70 ff.). Es gibt Varianzanalysen mit und ohne Messwiederholungen, beide Arten können sowohl univariat als auch multivariat sein (vgl. Bortz und Schuster 2010, S. 471 f.). Bei einer zweifaktoriellen Varianzanalyse mit Messwiederholung kann sich die Messwiederholung auf einen oder beide Faktoren beziehen. Wenn nur einer der beiden Faktoren ein Faktor der Messwiederholung ist, dient der andere Faktor dazu, die Testpersonen zu gruppieren (vgl. ebd., S. 288).

Da Varianzanalysen ebenso wie der t-Test zu parametrischen statistischen Verfahren zählen, ist die Intervallskalenqualität der abhängigen Variablen Grundvoraussetzung für die Anwendung dieser Verfahren. Daneben gibt es weitere Voraussetzungen, die erfüllt sein sollten. Für univariate Varianzanalysen zählen dazu die Normalverteilung der unabhängigen und abhängigen Variablen, die Varianzhomogenität und die Unabhängigkeit der Messwerte aller Bedingungen (vgl. Rasch et al. 2014, S. 48 f.). Die Voraussetzungen multivariater Varianzanalysen entsprechen denen der univariaten Varianzanalysen, wobei sich die Normalverteilungsannahme der abhängigen Variablen auf eine multivariate Normalverteilung erweitert und die Varianzhomogenität durch die Homogenität von Varianz- und Kovarianzmatrizen ersetzt wird

(vgl. Rudolf und Müller 2012, S. 120). Die Voraussetzung Unabhängigkeit der abhängigen Variablen in den einzelnen Faktorstufen wird nur bei Varianzanalysen ohne Messwiederholung erfüllt, weshalb die beiden Verfahren mit und ohne Messwiederholung in der Literatur gesondert behandelt werden. Bei Varianzanalysen mit mindestens zwei Messwiederholungen sollte als weitere Bedingung die Sphärizität erfüllt sein, die mithilfe des Mauchly Tests überprüft werden kann. Sphärizität bezeichnet, dass die Varianzen der Differenzen der Mittelwerte zweier Faktorstufen gleich sind. Ist Sphärizität nicht gegeben, kann die Greenhouse-Geisser Korrektur angewendet werden (vgl. Rasch et al. 2014, 71 f.). Nach Finch (2005) sind Varianzanalysen sehr robust gegenüber Verletzungen der Normalverteilung. Dies belegen auch Studien wie beispielsweise die Arbeit von Blanca et al. (2017). Auch wird davon ausgegangen, dass bei Verletzungen der Varianzhomogenität die Ergebnisse mit robusten Testverfahren dennoch zuverlässig sind (vgl. Rudolf und Müller 2012, S. 101). Idealerweise sollten ein mittlerer Stichprobenumfang und eine gleiche Versuchspersonenanzahl pro Faktorstufe erfüllt sein, damit Verletzungen der Voraussetzungen zu keinen Problemen führen (vgl. Rasch et al. 2014, S. 49).

Die Hauptvoraussetzung der Intervallskalenqualität ist in der hier vorliegenden Studie für alle abhängigen Variablen gegeben. Die Score-Variablen sind eindeutig intervallskaliert, da die Abstände der Daten genau definiert sind (vgl. Lamberti 2001, S. 34). Die Stufen-Variablen lassen sich nicht eindeutig zuordnen, werden aber als intervallskaliert gesehen. Auf den ersten Blick entsprechen sie einer Ordinalskala, da es eine Rangordnung zwischen den Merkmalen und ein Verhältnis in einer „Über-/Unterordnung“ (Früh 2007, S. 32) gibt. Normalerweise bleibt der Abstand der einzelnen Werte einer Rangordnung unklar, beispielsweise bei Siegerplatzierungen oder der Angabe von Gefühlsempfindungen. Doch in bestimmten Fällen können auch Rechnungen durchgeführt werden, weshalb man dann von einer quasi-metrischen Skala ausgeht. Brosius und Koschel (2012) fassen dies mit folgenden Worten zusammen: „[A]uch Skalen mit nur drei, vier oder fünf Skalenpunkten [werden] so behandelt, als ob die Abstände zwischen den einzelnen Skalenpunkten gleich groß wären“ (ebd., S.38). Als Folge dieser Annahme werden dann Berechnungen durchgeführt, die strenggenommen nicht zulässig sind. Schulnoten werden oftmals als solche, auch quasi-metrisch genannte Variablen behandelt (vgl. Hafner 2013, S. 36). Die Reflektierstufen, die sowohl für die Reflexions- als auch Reflexivitätsfähigkeit aus fünf Stufen und Zwischenstufen bestehen, werden in dieser Studie als intervallskaliert betrachtet, da die Intervalle zwischen den einzelnen Stufen als gleich aufgefasst werden, was nach Kroeber-Riel und Weinberg (1999, S. 193) zulässig ist. Bortz und Döring (2006, 181 f.) weisen darauf hin, dass dabei mindestens vier verschiedene Ausprägungen vorherrschen sollten und der berechnete Mittelwert sinnvoll interpretierbar sein sollte. Diese Voraussetzungen sind in dieser Studie gegeben, weshalb die Stufen-Variablen ebenfalls als intervallskaliert im Sinne einer quasi-metrischen Variable gesehen werden können und folglich

die Intervallskalenqualität erfüllt ist. Ein weiterer Aspekt, der das eben beschriebene Verständnis befürwortet, ist der, dass es für das komplexe Verfahren der multivariaten zweifaktoriellen Varianzanalyse mit Messwiederholung keine eindeutige Entsprechung im non-parametrischen Testverfahren gibt und dies nur über mehrere unterschiedliche Tests realisiert werden könnte. Die Normalverteilung der unabhängigen Variablen ist dadurch gewährleistet, dass die Variable kategorial ist. Die Gruppen werden zunächst aufgrund der vorhandenen Klassen gebildet, die Experimentalgruppe wird anschließend noch auf Grundlage erster Ergebnisse in Gruppen mit gleichen Merkmalen eingeteilt.

Die weiteren Voraussetzungen für Varianzanalysen können vor der Durchführung durch unterschiedliche Verfahren geprüft werden. Die Normalverteilung der abhängigen Variablen lässt sich durch visuelle Methoden oder durch statistische Tests, wie dem Kolmogorow Test oder den Shapiro-Wilk Test berechnen (vgl. Cleff 2019, S. 218 f.). Nach Meinung mehrerer Autoren hat der Shapiro-Wilk Test eine höhere statistische Power als der Kolmogorow Test, weshalb er für diese Studie verwendet wird (vgl. Mohd Razali und Yap 2011; Steinskog und Kvamstø 2007). Die Homogenität von Varianz- und Kovarianzmatrizen kann mithilfe des Levene Tests und des Box-Tests geprüft werden (vgl. Bühner und Ziegler 2009, S. 257). Sind die Varianzen ungleich, wird statt der herkömmlichen einfaktoriellen Varianzanalyse die Welch-Varianzanalyse bevorzugt, da sie gegenüber der Verletzung der Varianzhomogenität noch robuster ist (vgl. Büning 1991, S. 132). Die vierte Voraussetzung der Unabhängigkeit der Messwerte aller Bedingungen wird bei Varianzanalysen ohne Messwiederholung durch die zufällige Zuordnung der Testpersonen in Faktorstufen gewährleistet. Somit wird jede Versuchsperson nur einer konkreten Bedingung zugeteilt (vgl. Rasch et al. 2014, S. 49).

In der hier vorliegenden Studie finden zur Hypothesenprüfung multivariate zweifaktorielle Varianzanalysen mit Messwiederholung Anwendung. Dabei werden die Unterschiede der abhängigen Variablen zwischen den unterschiedlichen Gruppen bzw. Faktorstufen untersucht, als auch die Entwicklungen der abhängigen Variablen in Bezug auf Messzeitpunkte. Die unabhängigen Variablen werden in der nachfolgenden Auswertung als Faktor und deren Ausprägungen als Faktorstufen bezeichnet. Die abhängige Variable Reflektierfähigkeit besteht aus mehreren Teilvariablen. Deshalb findet stets ein multivariates Verfahren Anwendung. Eine multivariate zweifaktorielle Varianzanalyse setzt sich aus mehreren Teilauswertungen zusammen: Zunächst wird der Interaktionseffekt interpretiert, der besagt, ob Mittelwertsunterschiede auf die Kombination einzelner Faktoren zustande kommen (vgl. Kuckartz et al. 2013, S. 196). Ist dem so, werden anschließend die einfachen Haupteffekte des Zwischensubjektfaktors²⁸ und des Innersubjektfaktors²⁹ genauer untersucht und bei Signifikanzen post-hoc Tests durchgeführt, um konkretere Aussagen über vorhandene Unterschiede treffen zu können. Die

²⁸ Zwischensubjektfaktor ist in dieser Studie stets die Gruppe.

²⁹ Innersubjektfaktor ist in dieser Studie stets die Zeit bzw. der Messzeitpunkt.

Haupteffekte werden für die Variablen bestimmt, die keine Interaktionseffekte besitzen. Außerdem wird beim Innersubjektfaktor auf die Effektstärken nach Cohens d eingegangen, um die Wirkung der Intervention präzisieren zu können.

Die Varianzanalysen untersuchen folglich jeweils drei Aspekte, um die Hypothesen zu prüfen: Erstens werden die abhängigen Variablen zum ersten Messzeitpunkt der Faktorstufen miteinander verglichen, zweitens werden die entsprechenden Variablen der weiteren Messzeitpunkte zwischen den Gruppen verglichen und drittens werden die Differenzen der Variablen der Messzeitpunkte der einzelnen Gruppen untersucht und ihre Effektstärken berechnet. Die erste Untersuchung dient dazu, festzustellen, ob die Ausgangsbedingungen aller drei Faktorstufen gleich sind, sodass eine Vergleichbarkeit herrscht. Die zweite Untersuchung hat das Ziel, zu untersuchen, ob die Faktorstufen unterschiedliche Endergebnisse aufweisen, sodass die Veränderungen der abhängigen Variablen auf die unabhängigen Variablen zurückzuführen sind. Die dritte Untersuchung widmet sich genauer den einzelnen Veränderungen innerhalb der Gruppen. Durch berechnete Effektgrößen können Aussagen über die Wirkung der Intervention auf die jeweiligen Faktorstufen getroffen und diese gegenübergestellt werden.

3.4.3.2 Hypothesen und Vorgehen bei Auswertung

Folgende Hypothesen, die sich aus der Forschungsfrage ableiten, sollen mit Hilfe der statistischen Analyse geprüft werden:

Wissenschaftliche Hypothese 1:

Die Intervention bestehend aus zwei Teilinterventionen führt zu einer gesteigerten Reflektierfähigkeit.

Wissenschaftliche Hypothese 2:

Die Teilintervention in Form von Gibbs Reflexivitätszyklus führt zu einer Verbesserung der Reflexivitätsfähigkeit.

Wissenschaftliche Hypothese 3:

Die Teilintervention in Form der Hütemethode führt zu einer Verbesserung der Reflexionsfähigkeit.

Wissenschaftliche Hypothese 4:

Die Intervention führt zu einer Veränderung der Begriffsdefinition von Reflektieren.

Zur Testung der ersten Hypothese wird die Gruppe als Zwischensubjektfaktor mit den drei Faktorstufen Experimentalgruppe 1, Experimentalgruppe 2 und Kontrollgruppe definiert. Erstere erlernen die Reflektiermethoden in der Reihenfolge Gibbs Reflexivitätszyklus - Hütemethode, die zweite Gruppe erlernt sie in umgekehrter Reihenfolge und die Kontrollgruppe erhält keine der beiden Reflektiermethoden. Der Messzeitpunkt wird als Innersubjektfaktor definiert. Dieser besitzt die zwei Faktorstufen Messzeitpunkt 1 und Messzeitpunkt 2. Ersteres entspricht der Messung vor Durchführung der Intervention und umfasst die Werte der abhängigen Variablen des ersten Blogbeitrags zum Thema Tierschutz, der im Unterricht verfasst wird. Messzeitpunkt 2 entspricht der Messung nach der Intervention, ebenfalls in Form des letzten Blogbeitrags zum Thema Tierschutz, der auch im Unterricht geschrieben wird. Sowohl beide Teilgruppen der Experimentalgruppe als auch die Kontrollgruppe, die keine Intervention erfährt, verfassen die beiden Beiträge unter gleichen Bedingungen. Es handelt sich folglich um eine 3x2 Varianzanalyse mit zwei Faktoren und eine Messwiederholung auf dem zweiten Faktor. Die abhängige Variable Reflektierfähigkeit besteht dabei aus sieben Teilvariablen: Gesamtscore im Reflektieren, Reflexionsscore, maximal erreichte Reflexionsstufe, am häufigsten erreichte Reflexionsstufe, Reflexivitätsscore, maximal erreichte Reflexivitätsstufe und am häufigsten erreichte Reflexivitätsstufe.

Für die Überprüfung der zweiten und dritten Hypothese gibt es jeweils zwei Faktorstufen des Zwischensubjektfaktors Gruppe. Diese sind Experimentalgruppe 1 und 2. Der Innersubjektfaktor Zeitpunkt hat drei Messzeitpunkte: Messzeitpunkt 1 liegt vor der ersten Teilintervention und umfasst die Durchschnittswerte der abhängigen Variablen der ersten beiden Beiträge, die zuhause verfasst werden. Messzeitpunkt 2 liegt zwischen der ersten und zweiten Teilintervention und umschließt die Durchschnittswerte der abhängigen Variablen der dritten, vierten und fünften Beiträge, die ebenfalls zuhause geschrieben werden und Messzeitpunkt 3 findet nach der zweiten Teilintervention statt und umfasst die durchschnittlichen Variablenwerte des sechsten, siebten und achten Beitrags, die weiterhin zuhause geschrieben werden. Die abhängigen Variablen bestehen aus den beiden Hauptvariablen Reflexions- und Reflexivitätsfähigkeit, die sich aus mehreren einzelnen Variablen zusammensetzen. Die Reflexionsfähigkeit beinhaltet den Reflexionsscore, die maximal erreichte Reflexionsstufe und die am häufigsten erreichte Reflexionsstufe, die Reflexivitätsfähigkeit besteht aus dem Reflexivitätsscore, der maximal erreichten Reflexivitätsstufe und der am häufigsten erreichten Reflexivitätsstufe. Bei diesem Testverfahren handelt es sich um eine 2x3 Varianzanalyse mit zwei Faktoren und zwei Messwiederholungen auf dem zweiten Faktor.

Die vierte Hypothese wird analog zur ersten Hypothese getestet. Hierbei sind nur die abhängigen Variablen anders, da hier Bezug auf die Begriffsdefinition genommen wird und nicht auf die Entwicklung der Reflektierfähigkeiten. Die beiden abhängigen Variablen bei der Testung dieser Hypothese sind die fachliche Nähe und die enthaltenen Teilbereiche des Reflektierens.

Neben den vordefinierten Hypothesen werden weitere Hypothesen anhand der quantitativen Daten explorativ generiert. Diese untersuchen zum einen den Einfluss der Reihenfolge des Erhalts der Reflektierhilfen, des Geschlechts und der persönlichen Bewertung der Reflektiermethoden auf die Entwicklung der Reflektierfähigkeiten und zum anderen die Veränderung der persönlichen Einstellungen zum Reflektieren durch die Intervention. Für den Einfluss der Reihenfolge der Reflektierhilfen dienen die Ergebnisse von Hypothese 1, 2 und 3 als Grundlage für die weitere Untersuchung. Für die Untersuchung des Einflusses des Geschlechts hinsichtlich der Entwicklung der Reflektierfähigkeiten durch die Intervention werden der Zwischensubjektfaktor Geschlecht und der Innersubjektfaktor Messzeitpunkt 1 und 2 (erster und letzter Beitrag) gebildet. Bei der Analyse des Einflusses der persönlichen Bewertung der Reflektierhilfen auf die Entwicklung der Reflektierfähigkeiten wird der Zwischensubjektfaktor Bewertung *hilfreich* und Bewertung *nicht hilfreich* und der Innersubjektfaktor Messzeitpunkt 1 und 2 (vor und nach Intervention) definiert. Für die Hypothesengenerierung hinsichtlich der Entwicklung der persönlichen Einstellungen zum Reflektieren dient der Zwischensubjektfaktor Gruppe mit den beiden Experimentalgruppen und der Kontrollgruppe und der gleiche Innersubjektfaktor wie zuvor. Die abhängigen Variablen hierfür sind die Einschätzungen der persönlichen Einstellungen gegenüber dem Reflektieren, die über den Fragebogen mit einer Bewertung folgender Aussagen erhoben wurden (siehe Fragebogen im Anhang, S. A2):

- Ich treffe Entscheidungen eher spontan, ohne viel Nachzudenken.
- Ich habe immer eine Erklärung für mein Verhalten.
- Ich denke oft über mein Verhalten nach, um die Gründe dafür verstehen zu können.
- Ich akzeptiere andere Ansichten oder Meinungen bei komplexen Themen.
- Manchmal frage ich mich, wie andere etwas tun und versuche eine bessere Lösung zu finden.
- Ich denke oft in meiner Freizeit über Unterrichtsinhalte nach.

Die Ergebnisse der Hypothesentestungen und -generierungen sind in Kapitel 3.5.3 aufgeführt.

3.5 Ergebnisse

In diesem Kapitel werden die Ergebnisse beider Auswertungen vorgestellt. Zunächst wird auf die Einzelauswertungen der Probandinnen und Probanden im Rahmen der qualitativen Inhaltsanalyse eingegangen, anschließend erfolgt eine Zusammenfassung dieser Ergebnisse mit Einteilungen der Testpersonen in Gruppen hinsichtlich ihrer Entwicklungen der Reflektierfähigkeiten. Die Ergebnisdarstellung der quantitativen Auswertung findet in Form der Hypothesentestung und -generierung statt.

3.5.1 Einzelauswertungen der Teilnehmenden im Rahmen der qualitativen Inhaltsanalyse

Im Folgenden wird die Einzelauswertung der Testperson P5 dargestellt. Sie stellt exemplarisch das eben beschriebene, theoretische Vorgehen der Einzelanalysen aller Probadinnen und Probanden dar. Aus Gründen der Übersichtlichkeit sind weitere Einzelauswertungen im Anhang zu finden (vgl. Anhang, S. A16).

Einzelauswertung P5: – Allgemeines zur Person

Tabelle 14 Übersicht P5

Name	P5
Klasse	1
Geschlecht	Weiblich
Anzahl Beiträge	10 Beiträge
Reihenfolge Reflektiermethoden	Gibbs – Hüte
Verständnis von Reflektieren (vor Blogprojekt)	„Reflektieren bedeutet für mich, dass etwas oder jemand gespiegelt wird. Man kann aber auch Verhalten reflektieren bzw. kann Verhalten auch auf andere Menschen reflektiert werden.“
Verständnis von Reflektieren (nach Blogprojekt)	„Reflektieren bedeutet für mich sich mit einem Thema genau auseinander zusetzen [sic!] und es von verschiedenen Blickwinkeln zu betrachten. Dabei schaut man Vor- und Nachteile an und versucht sich mit Problematiken kritisch auseinanderzusetzen. Letztendlich bindet man auch sein eigenes Verhalten ein und hinterfragt es.“
Gibbs Reflexivitätszyklus Verständlichkeit	Stimmt eher schon
Gibbs Reflexivitätszyklus Hilfestellung	Stimmt eher schon
Hütemethode Verständlichkeit	Stimmt eher schon
Hütemethode Hilfestellung	Stimmt eher schon

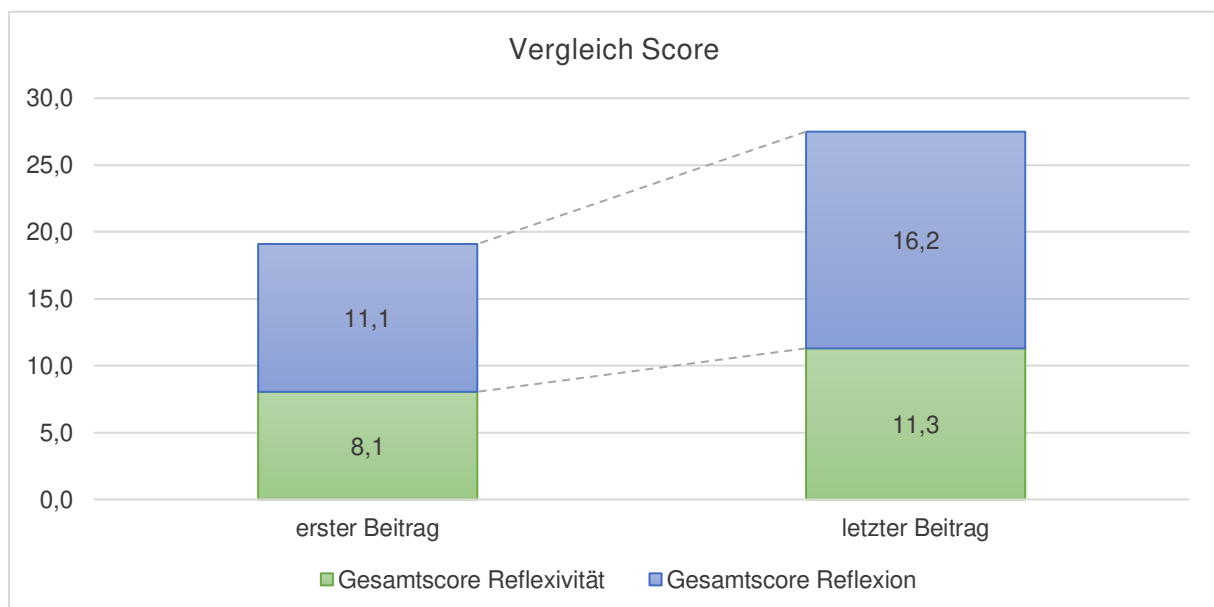
P5 ist weiblich und gehört zur Klasse 1. Insgesamt hat die Testperson 10 Beiträge verfasst. Nach den ersten drei Beiträgen erhält P5 die Reflexivitätshilfe, anschließend die Hütemethode. Zu Beginn des Projekts definiert P5 den Begriff *Reflektieren* mit folgenden Worten: „Reflektieren bedeutet für mich, dass etwas oder jemand gespiegelt wird. Man kann aber auch Verhalten reflektieren bzw. kann Verhalten auch auf andere Menschen reflektiert werden.“ (P5). Die Definition ändert sich am Ende des Projekts auf folgende: „Reflektieren bedeutet für mich sich mit einem Thema genau auseinander zusetzen [sic!] und es von verschiedenen Blickwinkeln zu

betrachten. Dabei schaut man Vor- und Nachteile an und versucht sich mit Problematiken kritisch auseinanderzusetzen. Letztendlich bindet man auch sein eigenes Verhalten ein und hinterfragt es.“ (P5). In der ersten Definition leitet sie den Begriff von der physikalischen Bedeutung des Spiegels ab, d.h. das gleiche Wiedergeben. Hier bezieht sie Reflektieren nur auf Reflexivität. Die zweite Definition beinhaltet eine multiperspektivische Sichtweise auf ein Thema und das Abwägen davon durch das Nennen von Vor- und Nachteilen. Während Reflexivität hier nicht mehr inkludiert ist, orientiert sich die Definition stark an dem Verständnis von Reflexion. P5 gibt im Fragebogen an, dass die Aussagen, beide Reflektierhilfen seien hilfreich und verständlich für sie eher schon zutreffen.

Vergleich der Beiträge zum Thema Tierschutz – Score

Der erste Beitrag von P5 weist einen Score im Reflektieren von 19,2 auf, der sich auf einen Reflexionsscore von 11,1 und einem Reflexivitätsscore von 8,1 verteilt (vgl. Abbildung 5). Der Score im Reflektieren des letzten Beitrags beträgt 27,5 und hat sich folglich um den Wert 8,4 gesteigert. Auch der Reflexionsscore des letzten Beitrags hat sich im Vergleich zum ersten Beitrag gesteigert (um 5,1 auf 16,2), ebenso der Reflexivitätsscore (um 3,2 auf 11,3). Sowohl die Reflexionsfähigkeit als auch die Reflexivitätsfähigkeit in Hinblick auf den Score haben sich vom ersten auf den letzten Beitrag von P5 zum Thema Tierschutz, der in der Klasse geschrieben wurde, leicht verbessert.

Abbildung 5 Scorewerte erster und letzter Beitrag



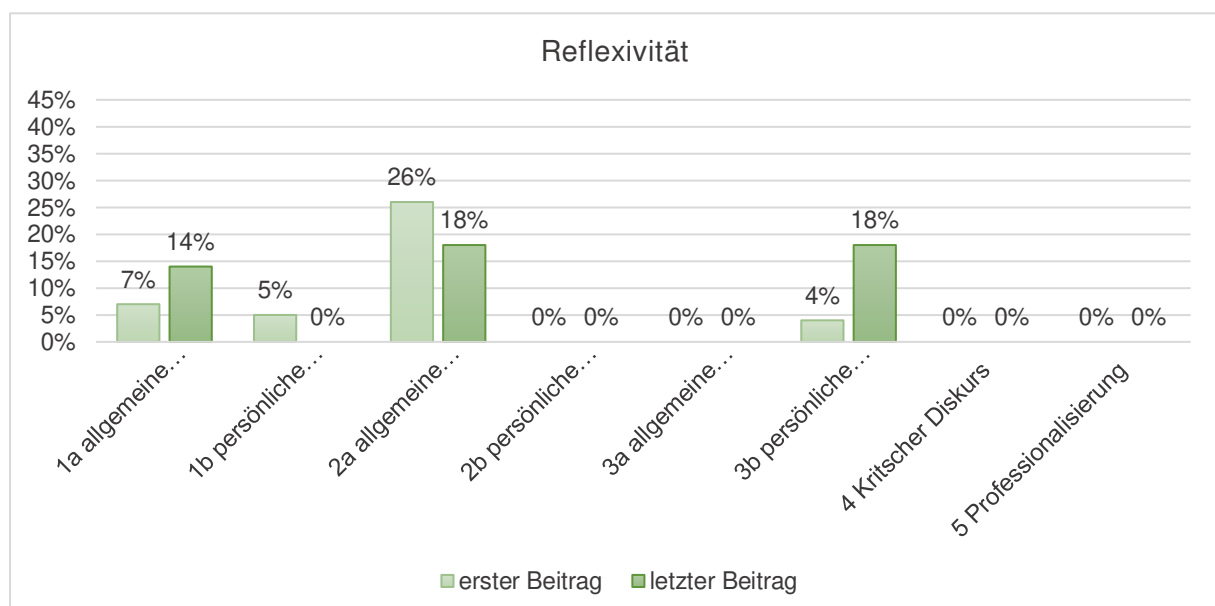
Vergleich der Beiträge zum Thema Tierschutz – Analyse der Reflektierstufen

Der erste Beitrag von P5 besteht zu 42% aus Textanteilen, die Reflexivitätsstufen zugeordnet werden können (vgl. Abbildung 6). Im ersten Beitrag erreicht P5 als höchste Stufe beim Reflektieren über die Handlung (Reflexivität) die Stufe 3b mit 4% Anteil am Gesamttext. Außerdem kommen die Stufen 1a (7%), 1b (5%) und 2a (26%) im ersten Beitrag vor. Im letzten Beitrag beträgt der Textanteil mit Reflexivitätsstufen 50%, die sich aus den Stufen 1a mit 14% und 2a und 3b mit jeweils 18% zusammensetzen. P5 reflektiert in beiden Beiträgen in den gleichen Stufen, jedoch mit einer unterschiedlichen Verteilung darauf. Im ersten Beitrag schreibt sie mehr über allgemeine Handlungsalternativen als über ihr eigenes Verhalten. Sie nennt einen zukünftigen Plan für ihr Verhalten, dass aus der Analyse des Themas selbst entwickelt wurde und geht zuvor nicht auf das momentane eigene Verhalten ein. Im zweiten Beitrag wird mehr auch über das eigene Verhalten reflektiert, indem sie über ihr aktuelles Verhalten nachdenkt und auch untersucht, welche Gefühle dabei auftreten:

Ich achte, seit ich begonnen habe mich mit dem Thema Nachhaltigkeit auseinander zusetzen [sic!], auch mehr darauf beim Einkaufen, dass die Eier von regionalen Hühnern aus Freilandhaltung und unser Fleisch möglichst nicht aus Massentierhaltung stammt. Dabei habe ich einfach ein besseres Gefühl. Dadurch kann man auf der einen Seite den Tieren ein bisschen helfen, man unterstützt Massentierhaltung nicht mehr und unterstützt regionale Betriebe, die sich für die Nachhaltigkeit einsetzen. Natürlich ist in diesem Bereich noch viel mehr möglich, aber aktuell ist es zumindest ein Versuch von mir, einen Schritt in die richtige Richtung zu machen.
(P5, letzter Beitrag, Reflexivitätsstufe 3b)

Obwohl die gleichen Reflexivitätsstufen im ersten und letzten Beitrag auftreten, reflektiert P5 im letzten Beitrag auf einem leicht höheren Niveau, da sie ausführlicher über ihr eigenes Verhalten reflektiert und nicht nur überwiegend über allgemeines Handeln schreibt.

Abbildung 6 Reflexivitätsverteilung erster und letzter Beitrag

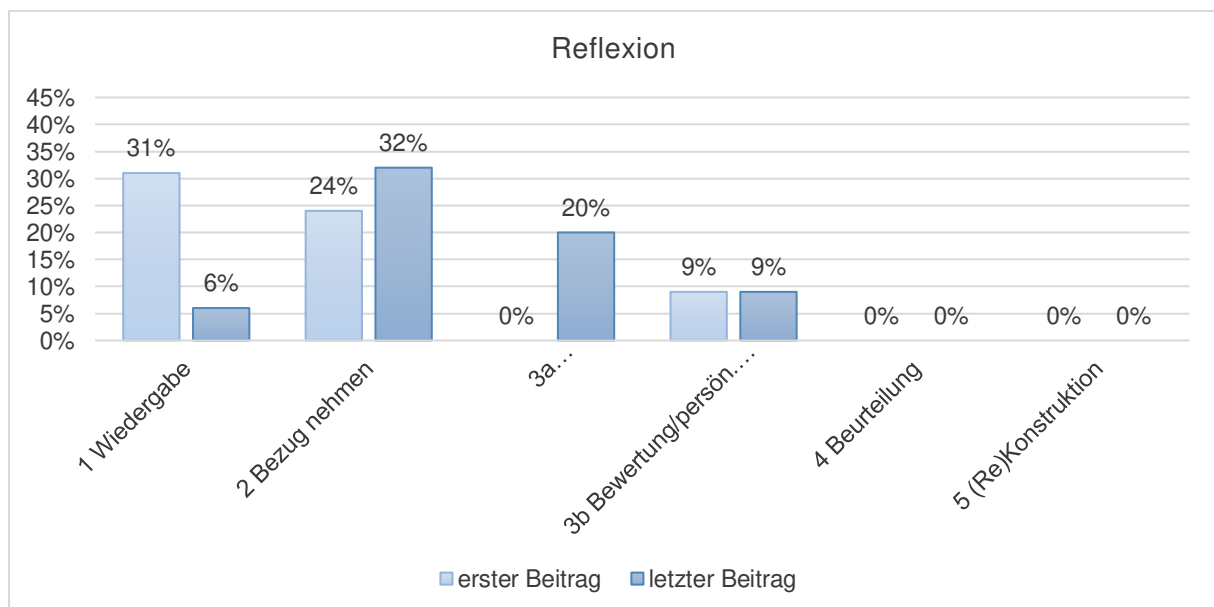


Im ersten Beitrag sind 64% des Textes mit Reflexionsanteilen (vgl. Abbildung 7). 31% davon zählen zu Stufe 1, 24% zu Stufe 2 und 9% zu Stufe 3b. P5 gibt überwiegend Fakten wieder und äußert ihre Meinung. In einem Satz gibt sie eine Begründung für ihre Einstellung an, weshalb sie Stufe 3b erreicht: „Ich persönlich muss zugeben, dass mich das Thema Tierschutz bisher nie besonders interessiert hat, da ich bis jetzt immer das Gefühl hatte, nichts dafür machen zu können“ (P5, erster Beitrag, Reflexionsstufe 3b). Der letzte Beitrag besteht zu 67% aus Reflexionsanteilen, ähnlich wie im ersten Beitrag. Diese verteilen sich auf die Stufen 1 (6%), 2 (32%), 3a (20%) und 3b (9%). Die Jugendliche nennt weniger reine Fakten als im ersten Beitrag, sondern nennt auch Gründe dafür, wie beispielsweise in folgendem Satz:

Und obwohl Tierschützer sich schon seit Jahrzehnten mit diesen Problemen auseinandersetzen und eifrig nach einer Lösung suchen, verenden dennoch jährlich immernoch ca. eine viertelmillion Meeresschildkröten durch die Langleinenfischerei, da sie von den Ködern angelockt werden. (P5, zehnter Beitrag, Reflexionsstufe 3a)

Auch nennt P5 Gründe für ihre eigene Meinung, weshalb sie wie im ersten Beitrag auch Stufe 3b als Maximalstufe erreicht. Die Reflexionsfähigkeit hat sich in Hinblick auf die maximal erreichte Stufe nicht verändert, jedoch kommt im letzten Beitrag Stufe 3a vor, die im ersten Beitrag nicht vorhanden ist. P5 reflektiert weniger oberflächlich über das Thema, sondern untersucht häufiger Gründe für Sachverhalte.

Abbildung 7 Reflexionsverteilung erster und letzter Beitrag



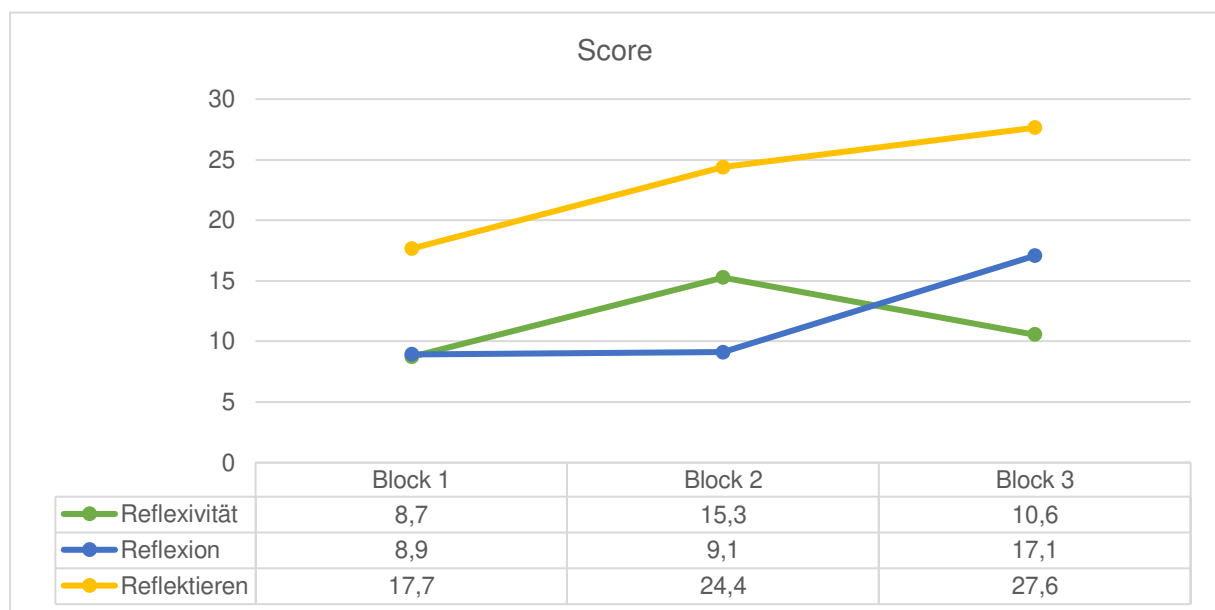
Vergleich der Beiträge vor, zwischen und nach den Reflektiermethoden – Score

Der Score im Reflektieren im ersten Textblock von P5 beträgt 17,7 und besteht aus einem Reflexivitätsscore von 8,7 und einem Reflexionsscore von 8,9 (vgl. Abbildung 8). Nach Erhalt

von Gibbs Reflexivitätszyklus steigt der Gesamtscore auf 24,4 an, auch der Reflexivitätsscore nimmt um einen ähnlichen Wert zu und erreicht 15,3. Der Reflexionsscore bleibt fast auf ähnlichem Niveau mit einem Wert von 9,1. Im dritten Textblock steigt der Gesamtscore nochmal um 3,2 auf 27,6 an. Der Reflexivitätsscore sinkt diesmal auf 10,6 und ist somit noch um etwa den Wert 2 höher als in Textblock an. Der Reflexionsscore steigt um 8 auf 17,1.

Zusammenfassend steht fest, dass der Gesamtscore durchgehend ansteigt, nach Gibbs Reflexivitätszyklus die Reflexivitätsfähigkeit in Hinblick auf den Score steigt und nach der Hütemethode die Reflexionsfähigkeit steigt.

Abbildung 8 Scorewerte der Textblöcke 1, 2 und 3



Vergleich der Beiträge vor, zwischen und nach den Reflektiermethoden – Analyse der Reflektierstufen

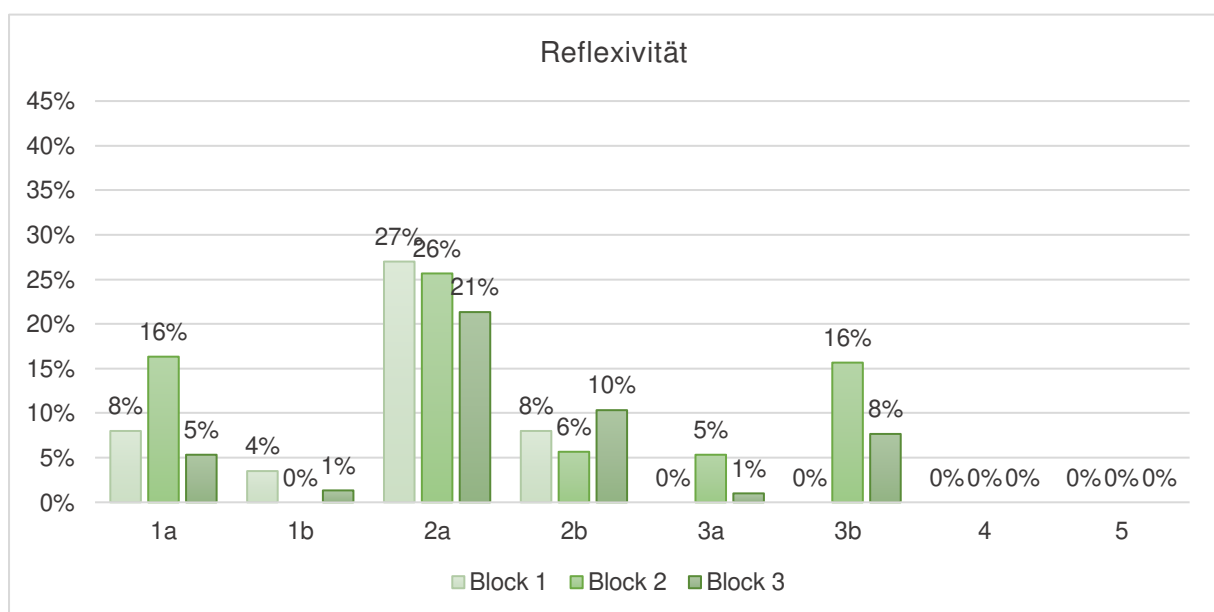
Im ersten Textblock besteht der Text zu 47% aus Reflexivitätsanteilen (vgl. Abbildung 9). Folgende Reflexivitätsstufen sind darin enthalten: 1a mit 8% Anteil am Gesamttext, 1b mit 4% Anteil, 2a mit 27% Anteil und 2b mit 8% Anteil. P5 reflektiert am häufigsten in Stufe 2a, d.h. sie begründet allgemeines Verhalten und schlägt mögliche Alternativhandlungen vor, wie beispielsweise im folgenden Satz: „Bei längeren Strecken sollte man möglichst öffentliche Verkehrsmittel nutzen. Immerhin ist der Bus das umweltfreundlichste Verkehrsmittel, gemessen am CO₂-Ausstoß pro Fahrgast und 100 Kilometer“ (P5, zweiter Beitrag, Reflexivitätsstufe 2a). Eigenes Verhalten begründet sie oberflächlich, ohne es genauer zu analysieren. Der zweite Textblock weist 69% Reflexivitätsanteile auf und somit 22% mehr als zuvor. Außerdem werden erstmals Stufe 3a und 3b mit 5% und 16% Anteil am Gesamttext erreicht. Stufe 1b ist nicht mehr Bestandteil im zweiten Textblock und Stufe 1a steigt um das doppelte auf 16% Anteil am

Gesamttext an. Die Stufen 2a und 2b sind in 26% und 6% des Textes vorhanden. Gibbs Reflexivitätszyklus, der darauf abzielt, die Reflexivitätsfähigkeit zu verbessern, hat bei P5 einen positiven Effekt auf die einzelnen Reflexivitätsstufen. Sowohl die maximal erreichte Stufe erhöht sich (von 2b auf 3b) und auch die Häufigkeit an Reflexivitätsstufen erhöht sich. Nach Gibbs Reflexivitätszyklus analysiert die Testperson nun die eigene Handlung und auch das allgemeine Verhalten und geht somit einen Schritt weiter als die reine Beschreibung und Begründung von Handlungsabläufen:

Wenn ich am Wochenende zusammen mit meinem Vater einkaufen gehe, versuchen wir meistens möglichst darauf zu achten, regionales Obst, Gemüse und Fleisch zu kaufen, weil man dabei einfach ein besseres Gefühl hat, da man weiß, wo es herkommt. Außerdem ist es umweltschonender und oftmals sogar billiger [als] das importierte Produkt. Natürlich könnte man auch immer die teuerste Alternative wählen. Aber teuer bedeutet nicht unbedingt, dass das Produkt eine bessere Qualität hat. Deshalb gibt es für mich eigentlich keine wirkliche Alternative und ich werde auch in Zukunft größtenteils regionale Produkte kaufen. (P5, vierter Beitrag, Reflexivitätsstufe 3b)

Im letzten Textblock sinkt die Häufigkeit von Reflexivitätsanteilen wieder auf 46% und liegt somit ein Prozent unter dem Anteil im ersten Textblock. Das Vorkommen aller Stufen nimmt vom zweiten auf den dritten Textblock ab, außer bei 2b, hier steigt sie um 4% auf 10% Anteil am Gesamttext an. Im dritten Textblock sind alle Stufen von 1a bis 3b enthalten (1a 5%, 1b 1%, 2a 21%, 2b 10%, 3a 1%, 3b 8%). Am häufigsten – wie auch bei den anderen beiden Textblöcken – kommt Stufe 2a vor. Somit bleibt die maximal erreichte Reflexivitätsstufe beim dritten Textblock gleich mit der maximal erreichten Reflexivitätsstufe beim zweiten Textblock, ihre Häufigkeit reduziert sich jedoch um etwa ein Drittel.

Abbildung 9 Reflexivitätsverteilung der Textblöcke 1, 2 und 3

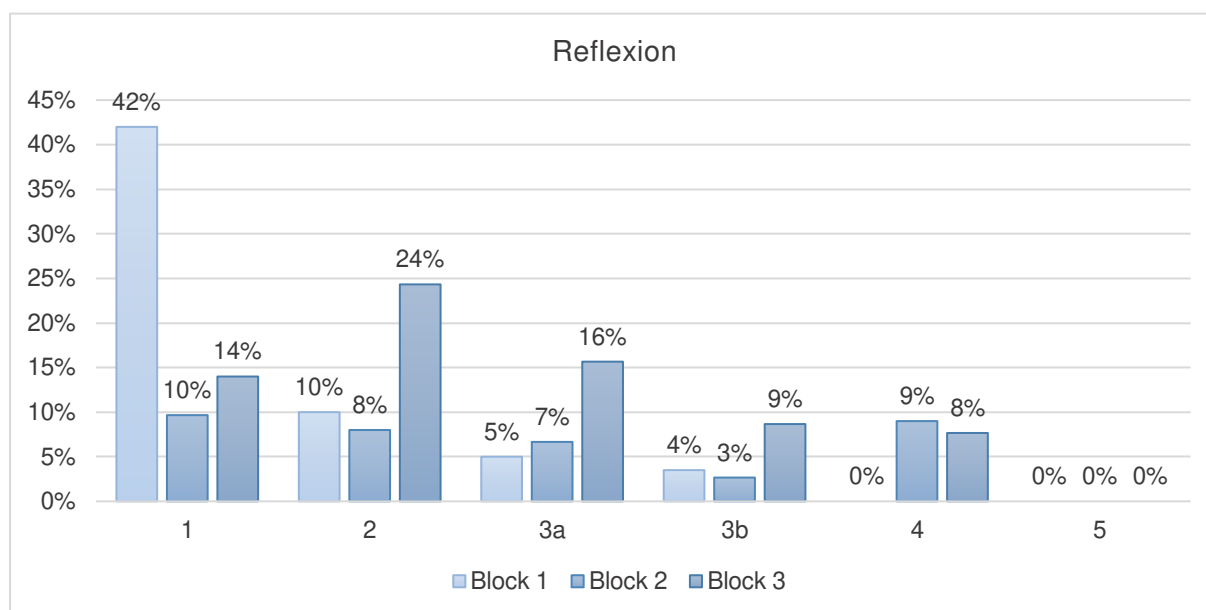


P5 weist im ersten Textblock und somit vor Erhalt der Reflektierhilfen folgende Reflexionsstufen auf: Stufe 1 mit 42% Anteil am Gesamttext, Stufe 2 mit 10% Anteil, Stufe 3a mit 5% Anteil und Stufe 3b mit 4% Anteil. Insgesamt besteht der Text von Block 1 zu 61% aus Reflexionsanteilen (vgl. Abbildung 10). Die Testperson gibt überwiegend Fakten wieder, ohne diese zu kommentieren. Folgendes Beispiel verdeutlicht die reine Wiedergabe von Inhalten: „Von 1850-2000 sind die Gletscherflächen der Alpen um 50% ihrer gesamten Fläche geschrumpft. Gleichzeitig steigt durch das rasante Schmelzen von Eis und Schnee auch der Meeresspiegel an“ (P5, dritter Beitrag, Reflexionsstufe 1). Im zweiten Textblock macht der Reflexionsanteil nur noch 37% des Gesamttextes aus und setzt sich aus den Stufen 1 (10%), 2 (8%), 3a (7%), 3b (3%) und 4 (9%) zusammen. Nach Erhalt der ersten Reflektierhilfe, in diesem Fall Gibbs Reflexivitätszyklus, reduziert sich die Häufigkeit von Stufe 1 um mehr als 30%, der prozentuale Anteil am Gesamttext der Stufen 2, 3a und 3b bleiben ungefähr gleich. Stufe 4 wird nach Gibbs erstmalig mit 9% Anteil am Text in einem Beitrag erreicht. Dabei untersucht die Testperson die Folgen von erhöhtem Onlinekaufen, diskutiert Probleme und sucht nach möglichen Lösungen. Stufe 4 in Textblock 2 kommt jedoch nur beim Thema Einkaufen vor, die anderen beiden weisen ebenfalls Stufe 3b als höchste erreichte Stufe auf. Der dritte Textblock, dessen Gesamttext zu 71% aus Reflexionsanteilen besteht, hat dieselben Reflexionsstufen wie der Textblock zuvor jedoch mit einer anderen Verteilung (14% Stufe 1, 24% Stufe 2, 16% Stufe 3a, 9% Stufe 3b, 8% Stufe 4). In diesem Textblock wird Stufe 4 in zwei von drei Beiträgen erreicht, die Stufen darunter kommen in allen drei Beiträgen vor. Folgender Textausschnitt zeigt eine tiefergehende Reflexion über den Inhalt inklusive Problemanalyse und Lösungsüberlegungen:

Perfekte Lösungen für das Problem der Stromgewinnungsfrage gibt es nicht. Alle Möglichkeiten haben ihre Nachteile, die den Vorteilen oftmals überwiegen. Aber keine Windräder oder auch keine Kohlekraftwerke zu bauen, ist auch unlogisch, da wir den Strom ja brauchen. Meiner Meinung nach ist Strom aus erneuerbaren Energien jedoch immer noch besser als Strom aus fossilen Energieträgern. Zumindest für meine eigene Gesundheit sind sie erheblich gesünder und durch gute Standortwahl kann ein Großteil der Nachteile (z.B. die Zerstörung von Lebensräumen) einfach wegfallen. Letztendlich würde es schon helfen, wenn mehr Leute Photovoltaikanlagen auf ihren Dächern hätten, da so jeder selber einen Teil seines Stroms gewinnen kann und die Gewährleistung hat, dass ihr Strom aus erneuerbaren Energien stammt. (P5, neunter Beitrag, Reflexionsstufe 4)

Im Vergleich der Textblöcke untereinander wird ersichtlich, dass die Häufigkeit von Reflexionsstufen im Gesamttext von Block 1 über Block 2 zu Block 3 stets ansteigt. Nach der zweiten Reflektierhilfe (Hütemethode) verändert sich die Verteilung der Reflexionsstufen deutlich: Stufe 2 kommt 16% häufiger vor, Stufe 3a 9% häufiger und Stufe 3b 6% häufiger. Der Anteil von Stufe 4 am Gesamttext reduziert sich minimal um 1%. Während nach Gibbs Reflexivitätszyklus Stufe 1 der Reflexion auffällig abfällt und Stufe 4 erstmalig vorkommt, bewirkt die Hütemethode eine Zunahme der Häufigkeiten der meisten Stufen, vor allem der Stufe 2 und 3a.

Abbildung 10 Reflexionsverteilung der Textblöcke 1, 2 und 3



3.5.2 Zusammenfassung der Ergebnisse der qualitativen Inhaltsanalyse

Im Folgenden werden die Ergebnisse der Einzelauswertung der Probandinnen und Probanden zusammengefasst dargestellt. Die Einzelauswertungen sind im Anhang aufgeführt (vgl. Anhang, S. A16). Die Daten entstammen der qualitativen Inhaltsanalyse, die am Material in Form der Blogbeiträge durchgeführt wurde. Zunächst wird ein Überblick über die Testpersonen gegeben, anschließend werden sie bezüglich der Veränderung ihrer Begriffsformulierung von *Reflektieren* nach der Intervention in Gruppen zusammengefasst. Im anschließenden Schritt werden die Teilnehmenden hinsichtlich ihrer Entwicklung der Reflektierfähigkeiten durch die Intervention in drei Gruppen eingeteilt und die Zusammensetzung dieser Gruppen detaillierter vorgestellt. Weiterhin werden die Testpersonen hinsichtlich ihrer Entwicklungen über die beiden Teilinterventionen hinweg je nach untersuchten Kriterium in Gruppen eingeteilt und zusätzlich Gruppen in Hinblick auf die Wirkung der ersten Teilintervention (Gibbs Reflexivitätszyklus oder Hütemethode) gebildet.

(1) Allgemeines zu Teilnehmenden

Die 24 Teilnehmenden aus zwei Schulklassen setzen sich insgesamt aus 13 männlichen und elf weiblichen Personen zusammen (vgl. Tabelle 15). Jeweils zwölf Testpersonen erhalten als erste Teilintervention Gibbs Reflexivitätszyklus bzw. die Hütemethode. P12 war zum Zeitpunkt des Erhalts der ersten Teilintervention nicht anwesend.

Tabelle 15 Übersicht Testpersonen

Testperson	Geschlecht	Klasse	Erste Teilintervention
P1	m	1	Hütemethode
P2	m	1	Gibbs Reflexivitätszyklus
P3	m	1	Hütemethode
P4	m	1	Gibbs Reflexivitätszyklus
P5	w	1	Gibbs Reflexivitätszyklus
P6	w	1	Gibbs Reflexivitätszyklus
P7	w	1	Gibbs Reflexivitätszyklus
P8	w	1	Hütemethode
P9	w	1	Hütemethode
P10	m	1	Hütemethode
P11	w	1	Hütemethode
P12	m	1	(Gibbs Reflexivitätszyklus) / krank
P13	m	2	Hütemethode
P14	m	2	Gibbs Reflexivitätszyklus
P15	m	2	Gibbs Reflexivitätszyklus
P16	w	2	Hütemethode
P17	w	2	Gibbs Reflexivitätszyklus
P18	w	2	Hütemethode
P19	m	2	Hütemethode
P20	m	2	Hütemethode
P21	w	2	Gibbs Reflexivitätszyklus
P22	w	2	Hütemethode
P23	m	2	Gibbs Reflexivitätszyklus
P24	m	2	Gibbs Reflexivitätszyklus

(2) Veränderung der eigenen Definition des Begriffs *Reflektieren*

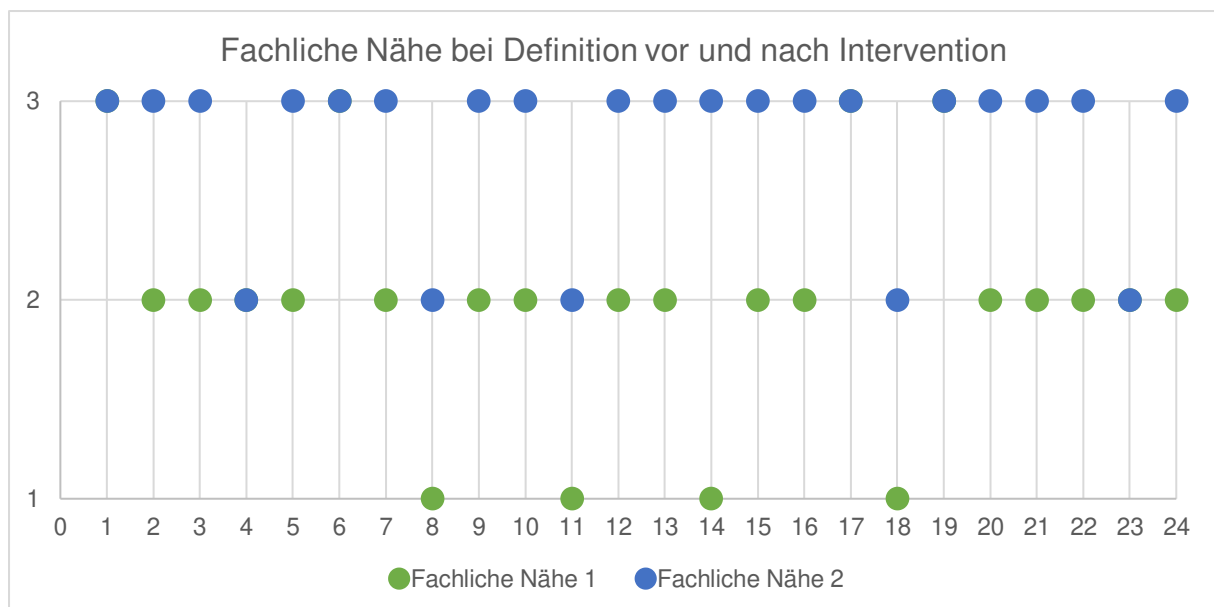
Die Schülerinnen und Schüler unterscheiden sich hinsichtlich ihrer Veränderung der eigenen Begriffsdefinition von *Reflektieren*. Für die Einteilung in Gruppen werden die Kriterien der fachlichen Nähe und der enthaltenen Teilbereiche des Reflektierens innerhalb der Definition herangezogen. Im Folgenden werden die zusammengefassten Ergebnisse vorgestellt, die einzelnen Definitionsformulierungen können in den Einzelauswertungen im Anhang nachgelesen werden.

Veränderung der fachlichen Nähe innerhalb der Definition

Vor der Intervention weisen die Testpersonen P8, P11, P14 und P18 keine fachliche Nähe in ihrer Definition zum Begriff *Reflektieren* auf (vgl. Abbildung 11). Das bedeutet, dass sie unter

dem Begriff kein Reflektieren im Sinne von Nachdenken verstehen, sondern den Begriff einem anderen Fachbereich, beispielsweise der Physik, zuordnen. 16 Teilnehmende ordnen den Begriff mit einer etwas fachlichen Nähe ein, d.h. sie integrieren oberflächlich ein Charaktermerkmal von Reflektieren. Ebenfalls vier Testpersonen (P1, P6, P17, P19) zeigen bereits vor der Intervention eine hohe fachliche Nähe in ihrer Definition. Die genauen Wortlaute der Definitionen können aus den Einzelanalysen im Anhang entnommen werden. Nach der Intervention zeigen die meisten Schülerinnen und Schüler eine höhere fachliche Nähe als zuvor, P4 und P23 bleiben auf dem gleichen Niveau von etwas fachlicher Nähe. Keine Testperson erreicht weniger fachliche Nähe als zuvor. P8, P11 und P18 steigern sich von keiner fachlichen Nähe auf etwas fachliche Nähe bei ihrer Definition des Begriffs *Reflektieren* und P14 erreicht eine hohe fachliche Nähe von zuvor keiner fachlichen Nähe. Die restlichen Testpersonen steigern sich von etwas fachlicher Nähe zu hoher fachlicher Nähe.

Abbildung 11 Fachliche Nähe bei Definition vor und nach Intervention



Hinsichtlich der fachlichen Nähe bei den eigenen Definitionen können die Teilnehmenden in fünf Gruppen eingeteilt werden, je nachdem welche Ausprägungen an fachlicher Nähe sie vor und nach der Intervention in ihren Definitionen aufweisen:

Tabelle 16 Übersicht Gruppeneinteilung in Hinblick auf fachliche Nähe bei Definition

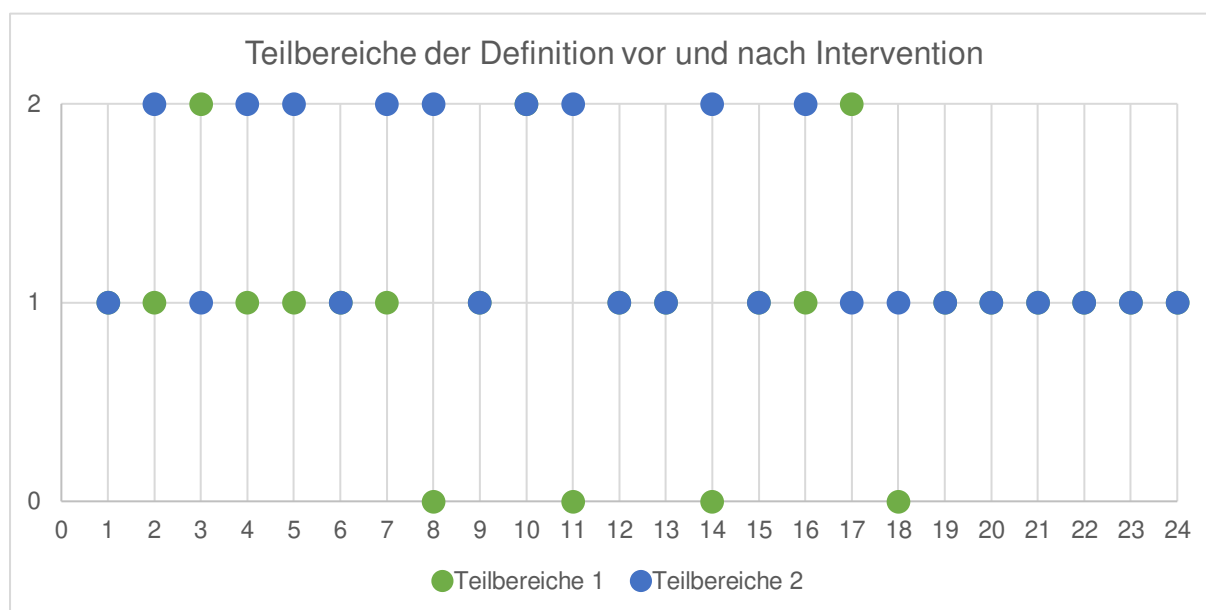
Gruppe 1 (keine fachliche Nähe vor Intervention, hohe fachliche Nähe nach Intervention)	<u>P14</u> ³⁰
Gruppe 2 (keine fachliche Nähe vor Intervention, etwas fachliche Nähe nach Intervention)	P8, P11, P18
Gruppe 3 (etwas fachliche Nähe vor Intervention, hohe fachliche Nähe nach Intervention)	<u>P2</u> , P3, <u>P5</u> , <u>P7</u> , P9, P10, P12, P13, <u>P15</u> , P16, <u>P20</u> , P21, P22, <u>P24</u>
Gruppe 4 (hohe fachliche Nähe vor und nach Intervention)	P1, <u>P6</u> , <u>P17</u> , P19
Gruppe 5 (etwas fachliche Nähe vor und nach Intervention)	<u>P4</u> , <u>P23</u>

Enthaltene Teilbereiche des Reflektierens

Vor der Intervention gibt es vier Testpersonen (P8, P11, P14, P18), die weder das Verständnis von Reflexion noch Reflexivität in ihrer Definition des Begriffs *Reflektieren* aufnehmen (vgl. Abbildung 12). Die Definitionen von P3 und P17 umfassen beide Teilbereiche des Reflektierens, die übrigen Probandinnen und Probanden integrieren einen der beiden Teilbereiche (P5 und P12 weisen den Teilbereich Reflexivität auf, die übrigen den Teilbereich Reflexion). Nach der Intervention besitzen alle Definitionen mindestens einen Teilaspekt (P3 und P9 Reflexivität, die übrigen Reflexion), P2, P4, P5, P7, P8, P10, P11, P14 und P16 weisen beide Teilaspekte auf. P3 und P17, die zuvor beide Teilbereiche aufgeführt haben, berücksichtigen in der zweiten Definition nur noch einen Teilbereich, die übrigen Testpersonen verbessern sich diesbezüglich. P8, P11 und P14 erwähnen in ihrer zweiten Definition sowohl den Aspekt der Reflexion als auch den der Reflexivität, zuvor haben sie keinen von beiden in der Definition aufgeführt. Die exakten Formulierungen der Definitionen können aus den Einzelanalysen im Anhang entnommen werden.

³⁰ Die unterstrichenen Testpersonen sind diejenigen, die als erste Reflektierhilfe Gibbs Reflexivitätszyklus erhalten, die ohne Unterstreichung diejenigen, die zunächst die Hütemethode von de Bono erhalten.

Abbildung 12 Teilbereiche des Reflektierens in Definition vor und nach Intervention



Insgesamt lassen sich sechs Gruppen hinsichtlich der Entwicklung der enthaltenen Reflektier-teilbereiche innerhalb der Definitionen vor und nach der Intervention ableiten:

Tabelle 17 Übersicht Gruppeneinteilung in Hinblick auf Teilbereiche des Reflektierens in Definition

Gruppe 1 (keinen Teilbereich vor Intervention, beide Teilbereiche nach Intervention)	P8, P11, <u>P14</u>
Gruppe 2 (keinen Teilbereich vor Intervention, einen Teilbereich nach Intervention)	P18
Gruppe 3 (einen Teilbereich vor Intervention, beide Teilbereiche nach Intervention)	<u>P2</u> , <u>P4</u> , <u>P5</u> , <u>P7</u> , P16
Gruppe 4 (einen Teilbereich vor und nach Intervention)	P1, <u>P6</u> , P9, P12, P13, <u>P15</u> , P19, <u>P20</u> , P21, P22, <u>P23</u> , <u>P24</u>
Gruppe 5 (beide Teilbereiche vor und nach Intervention)	P10
Gruppe 6 (beide Teilbereiche vor Intervention, einen Teilbereich nach Intervention)	P3, <u>P17</u>

Hinsichtlich der Veränderungen der Begriffsdefinitionen von Reflektieren kann festgehalten werden, dass die meisten Testpersonen Verbesserungen hinsichtlich der fachlichen Nähe und Zunahmen der enthaltenen Teilbereiche des Reflektierens zeigen. Die Intervention scheint

sich folglich positiv auf das individuelle Verständnis von Reflektieren auszuwirken. Ob diese Entwicklung tatsächlich durch die Intervention hervorgerufen wird, wird im späteren Verlauf im Rahmen einer nachgelagerten quantitativen Auswertung noch genauer untersucht (vgl. Kapitel 3.5.3).

(3) Entwicklung der allgemeinen Reflektierfähigkeiten vom ersten zum letzten Beitrag

Die Teilnehmenden lassen sich hinsichtlich der Entwicklung der allgemeinen Reflektierfähigkeit vom ersten zum letzten Beitrag in drei Gruppen einteilen: Testpersonen, die bei allen ausgewerteten Daten (drei Scores, zwei maximal erreichte Stufen und zwei am häufigsten erreichte Stufen) positive Entwicklungen zeigen, Schülerinnen und Schüler, die bei den ausgewerteten Daten sowohl positive als auch keine Entwicklungen zeigen und diejenigen, die bei den ausgewerteten Daten neben positiven oder keinen Entwicklungen auch negative Entwicklungen aufzeigen. Die nachfolgende Tabelle gibt einen Überblick über die Einteilung der Testpersonen in diese drei Gruppen:

Tabelle 18 Übersicht Gruppeneinteilung in Hinblick auf Entwicklung der Reflektierfähigkeit vom ersten zum letzten Beitrag

Gruppe 1 (nur positive Entwicklungen)	P3, <u>P6</u> , <u>P7</u> , P8, P13, <u>P14</u> , P21	
Gruppe 2 (positive und keine Entwicklungen)	1x keine Entwicklung	<u>P2</u> , P10, <u>P17</u> , P18, <u>P20</u> , P22, <u>P23</u>
	2x keine Entwicklung	<u>P5</u> , P11, <u>P15</u>
	3x keine Entwicklung	P19
	5x keine Entwicklung	P12
Gruppe 3 (auch negative Entwicklungen)	P1, <u>P4</u> , P9, P16, <u>P24</u>	

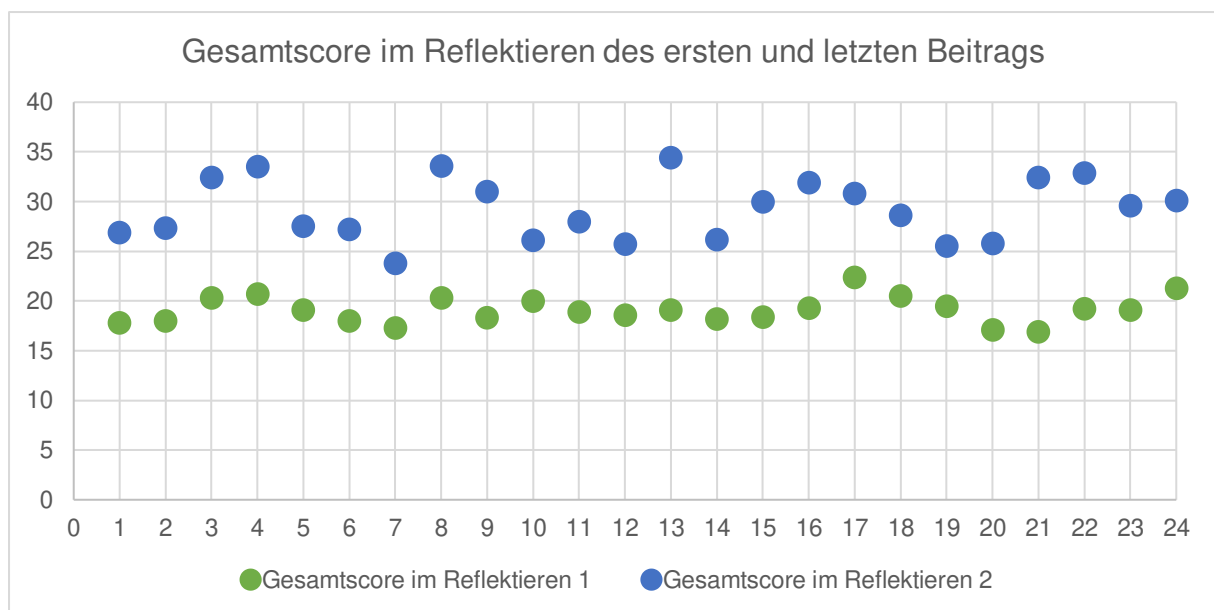
Die zweite Gruppe lässt sich noch in weitere Untergruppen aufteilen. Die Testpersonen P2, P10, P17, P18, P20, P22 und P23 besitzen einen neutral bleibenden Wert, während sich alle anderen Werte positiv verändern und bilden somit eine Untergruppe innerhalb der Gruppe 2. P5, P11 und P15 bilden eine weitere Untergruppe aus denjenigen, die zwei konstant bleibende Werte neben sich positiv verändernden Werten besitzen und P19 formt die dritte Untergruppe, da er drei unveränderte Werte aufweist. P12, der bei Erhalt der ersten Teilintervention krank war, hat fünf gleichbleibende Werte und bildet alleine die vierte Untergruppe (vgl. Tabelle 18). Bei Gruppe 3 weist P4 dreimal negative Entwicklungen auf, die übrigen drei Testpersonen besitzen jeweils nur eine negative Entwicklung. Im Folgenden werden die Entwicklungen der

einzelnen ausgewerteten Kriterien aller Probandinnen und Probanden zusammenfassend vorgestellt, damit nachvollzogen werden kann, wie sich die drei Hauptgruppen zusammensetzen.

Entwicklung des Gesamtscores im Reflektieren

Zum ersten Messzeitpunkt im Rahmen des ersten Blogbeitrags zum Thema Tierschutz, der im Unterricht geschrieben wird, variieren die Werte des Gesamtscores im Reflektieren zwischen 16,9 und 22,4 innerhalb der Experimentalgruppe (vgl. Abbildung 13). P21 weist dabei den niedrigsten Gesamtscore im Reflektieren auf und P17 den höchsten. P21 zeigt auch die größte Steigerung des Gesamtscores im Reflektieren vom ersten zum letzten Blogbeitrag mit einer Zunahme von 15,5, sodass der Wert nach der Intervention fast doppelt so hoch als zuvor ist. P19 hat die geringste Zunahme mit einem Wert von 6,0 (Messzeitpunkt 1: 19,5; Messzeitpunkt 2: 25,5). Den höchsten Gesamtscore im Reflektieren im letzten Beitrag besitzt P13 mit einem Wert von 34,4, den niedrigsten P7 mit 23,8.

Abbildung 13 Gesamtscore im Reflektieren – erster und letzter Beitrag



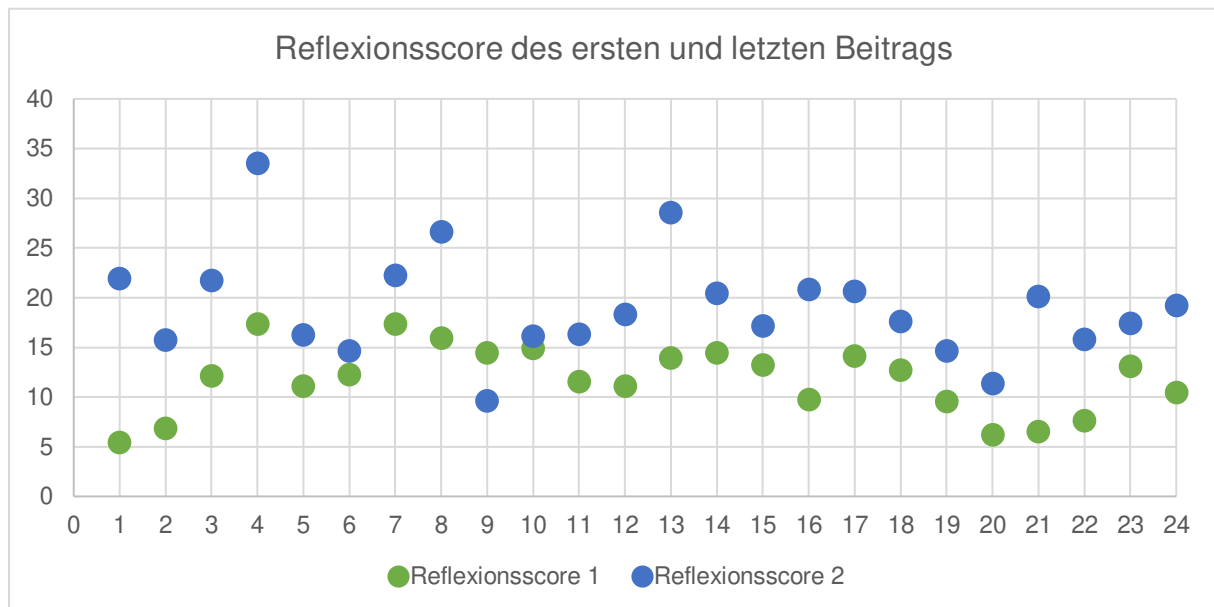
Alle Testpersonen verbessern sich hinsichtlich des Gesamtscores im Reflektieren von Messzeitpunkt 1 zu Messzeitpunkt 2 und haben folglich bessere Reflektierscorewerte als vor der Intervention. Alle Schülerinnen und Schüler sind beim ersten Beitrag auf einem ähnlichen Niveau, beim letzten Beitrag variieren die Werte des Gesamtscores im Reflektieren stärker.

Entwicklung des Reflexionsscores

Der Reflexionsscore des ersten Beitrags der einzelnen Testpersonen variiert zwischen den Werten 5,4 bei P1 und 17,3 bei P3 und P7 (vgl. Abbildung 14). Beim letzten Beitrag schwanken

die erreichten Reflexionsscorewerte zwischen 9,6 bei P9 und 33,5 bei P4. Die maximale Steigerung des Reflexionsscore hat dabei P1 mit einer Zunahme von 16,5. P9 entwickelt sich hinsichtlich des Reflexionsscores negativ mit einer Abnahme von - 4,8 von 14,4 auf 9,6. P10 bleibt von 14,9 beim ersten Beitrag auf 16.1 beim letzten Beitrag mit einer Differenz von 1,2 auf einem ähnlichen Reflexionsscore-Niveau.

Abbildung 14 Reflexionsscore – erster und letzter Beitrag

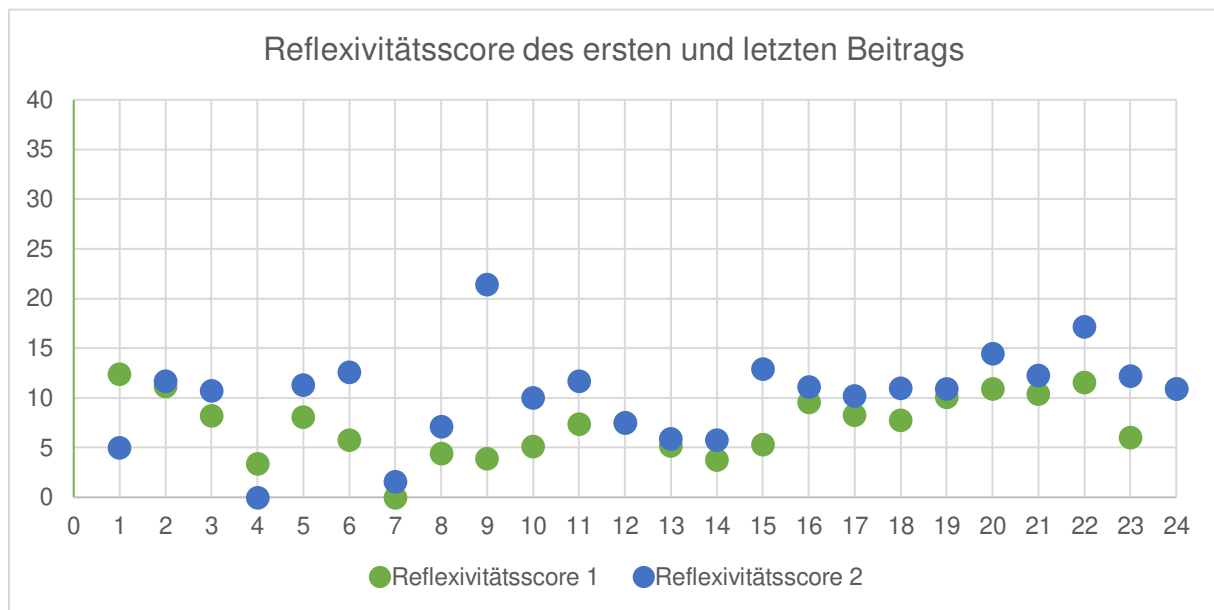


Alle Teilnehmenden mit Ausnahme von P9 haben vor der Intervention geringere Reflexionsscores als nach der Intervention. Manche Testpersonen zeigen am Ende deutlich höhere Werte als zuvor, die meisten nehmen um einen Wert zwischen 5 und 10 zu.

Entwicklung des Reflexivitätsscores

Ähnlich zum Reflexionsscore des ersten und letzten Beitrags ist auch die Verteilung des Reflexivitätsscores. Die Werte beim ersten Beitrag liegen mit Ausnahme einer Testperson alle unter den Werten des letzten Beitrags (vgl. Abbildung 15). Während der Reflexivitätsscore des ersten Beitrags zwischen Werten von 0,0 bei P7 und 12,4 bei P1 schwankt, bewegt er sich beim letzten Beitrag um Werte von 0,0 bei P4 und 21,4 bei P9. Die größte Veränderung zeigt dabei P9 mit einer Zunahme von 17,5. P1 hat die größte negative Veränderung mit einer Abnahme von - 7,4 von 12,4 auf 5,0. P12 und P24 zeigen keine Veränderung ihres Reflexivitätsscores vor und nach der Intervention.

Abbildung 15 Reflexivitätsscore – erster und letzter Beitrag

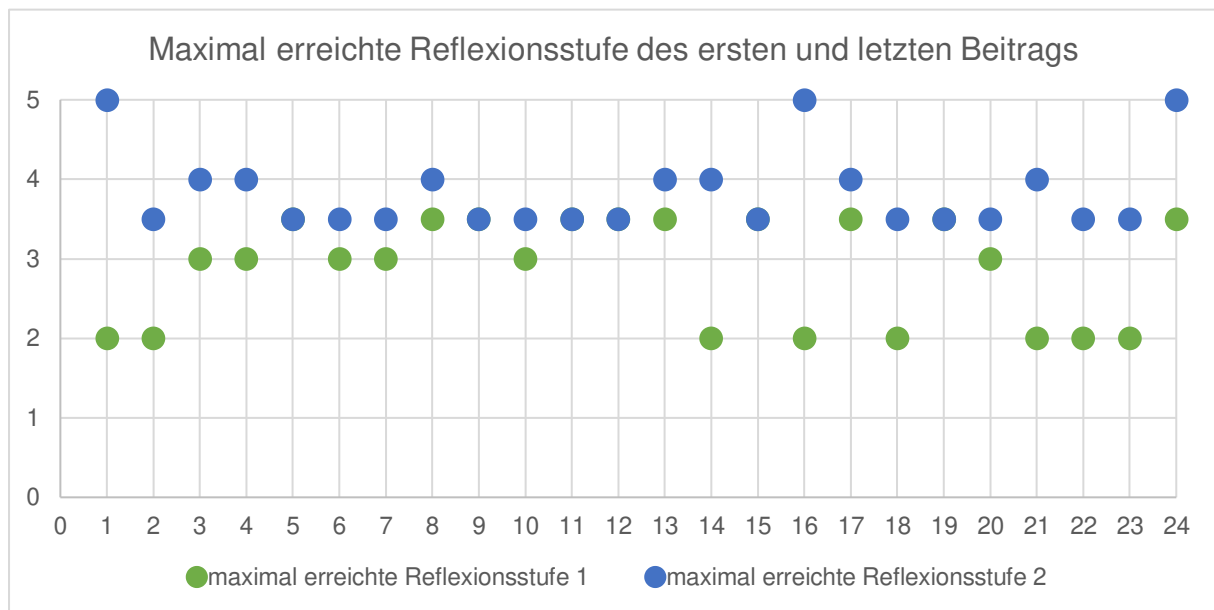


Alle Schülerinnen und Schüler mit Ausnahme von P4 zeigen keine negativen Entwicklungen des Reflexivitätsscores vom ersten zum letzten Blogbeitrag. Die meisten Testpersonen bessern sich diesbezüglich, zwei haben weiterhin den gleichen Reflexivitätsscore.

Entwicklung der maximal erreichten Reflexionsstufe

Die maximal erreichte Reflexionsstufe der Testpersonen des ersten Beitrags variiert zwischen der Stufe 2 und 3,5 (vgl. Abbildung 16). Die Probandinnen und Probanden P1, P2, P14, P16, P18, P21, P22 und P23 besitzen als maximal erreichte Reflexionsstufe Stufe 2, die übrigen liegen darüber. Im letzten Blogbeitrag liegen die Stufen bei allen Teilnehmenden gleich oder höher als zuvor. Keiner der Testpersonen erreicht nach der Intervention eine niedrigere Stufe als zuvor. P1, P16 und P24 erreichen als maximale Reflexionsstufe Stufe 5, die übrigen Probandinnen und Probanden eine Stufe zwischen 3 und 4. Die Testpersonen P5, P9, P11, P12 und P19 zeigen keine Änderung hinsichtlich der maximal erreichten Reflexionsstufe vom ersten zum letzten Beitrag. P1 und P16 haben mit einer Differenz von 3,0 die höchsten Veränderungen der maximal erreichten Reflexionsstufe.

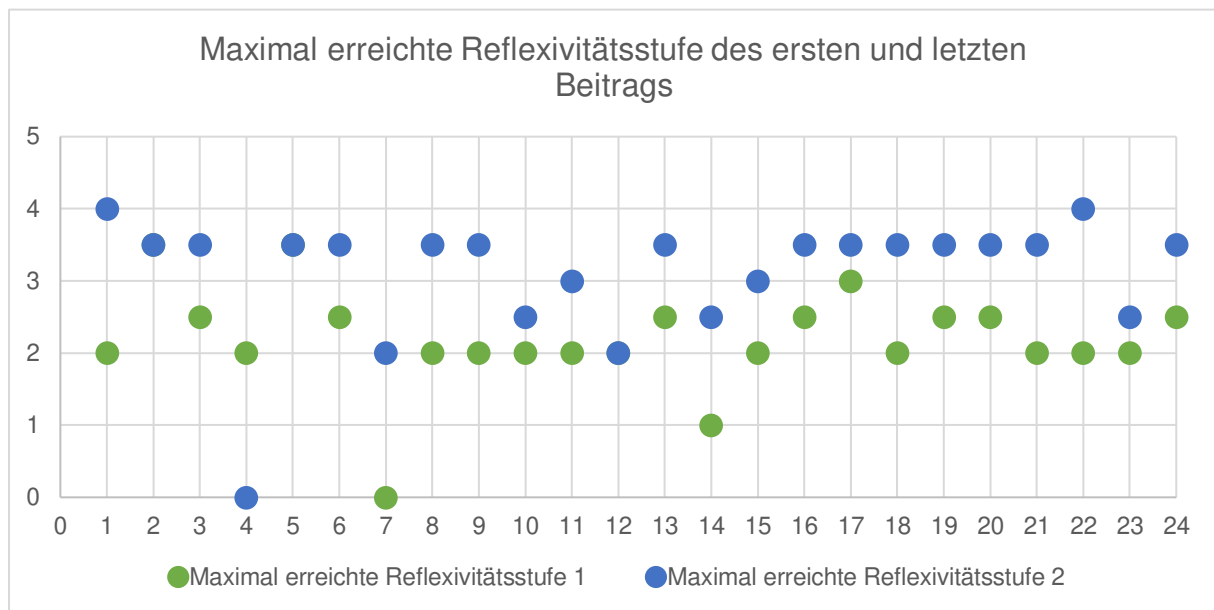
Abbildung 16 Maximal erreichte Reflexionsstufe – erster und letzter Beitrag



Entwicklung der maximal erreichten Reflexivitätsstufe

Die maximal erreichten Reflexivitätsstufen im ersten Beitrag liegen zwischen Null bei P7 und 3,5 bei P2 und P5 (vgl. Abbildung 17). Am häufigsten wird die Stufe 2 bei insgesamt elf Teilnehmenden erreicht. Im letzten Blogbeitrag erreichen die Testpersonen Werte zwischen 0 bei P4 und 4,0 bei P1 und P22 bei der maximal erreichten Reflexivitätsstufe. Die Stufe 3,5 wird dabei von 14 Jugendlichen erreicht und bildet somit die am häufigsten erreichte Maximalstufe. Alle Testpersonen mit Ausnahme von P2, P4, P5 und P12 haben im letzten Beitrag eine höhere maximal erreichte Reflexivitätsstufe erreicht als im ersten Beitrag. P4 zeigt hingegen einen Abfall der Stufe um zwei volle Stufen und erreicht keine Reflexivitätsstufe im letzten Beitrag. P2, P5 und P12 verändern sich hinsichtlich ihrer maximal erreichten Reflexivitätsstufe nicht. Die maximalen Zunahmen an Stufen zeigen sich bei P1, P7 und P22 mit einer Differenz von 2,0 vom ersten zum letzten Beitrag.

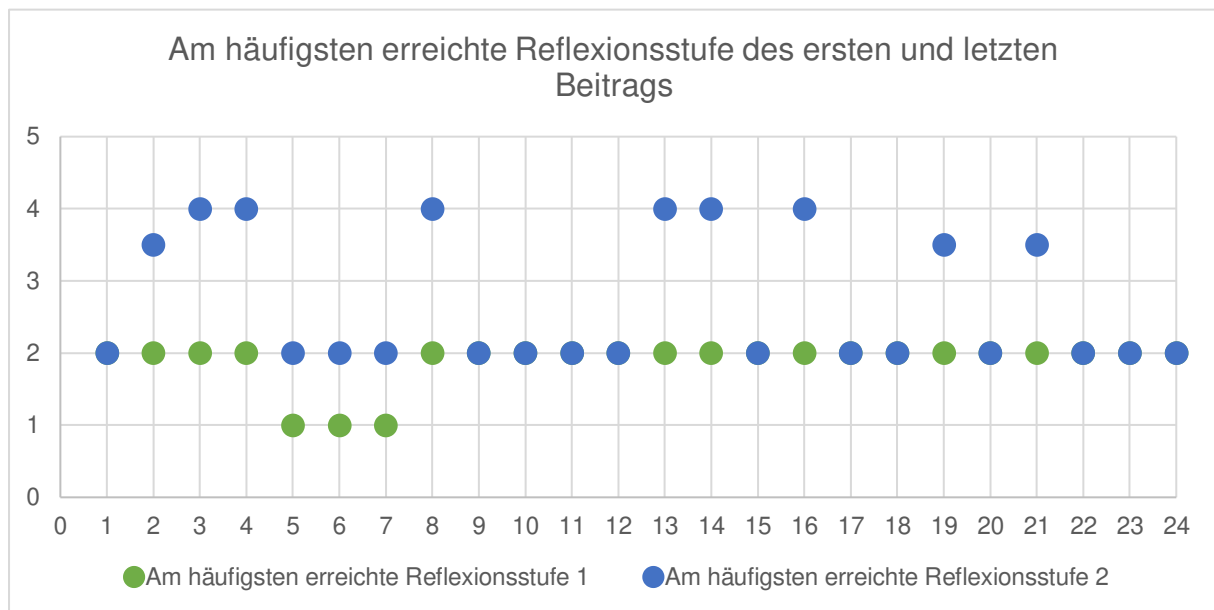
Abbildung 17 Maximal erreichte Reflexivitätsstufe – erster und letzter Beitrag



Entwicklung der am häufigsten erreichten Reflexionsstufe

Alle Testpersonen mit Ausnahme von P5, P6 und P7 haben Stufe 2 als am häufigsten erreichte Reflexionsstufe im ersten Beitrag (vgl. Abbildung 18). Die anderen drei erreichen nur Stufe 1. Im letzten Beitrag liegt die am häufigsten erreichte Reflexionsstufe bei Stufe 2 oder 4 bei allen Teilnehmenden mit Ausnahme von P2, hier ist es die Stufe 3,5. Die maximale Zunahme an Stufen beträgt den Wert 2,0 und wird von sechs Probandinnen und Probanden erreicht. Zwölf Testpersonen haben beim letzten Beitrag die gleiche am häufigsten erreichte Reflexionsstufe als zuvor. Keiner hat im letzten Beitrag eine niedrigere Stufe als im ersten Beitrag.

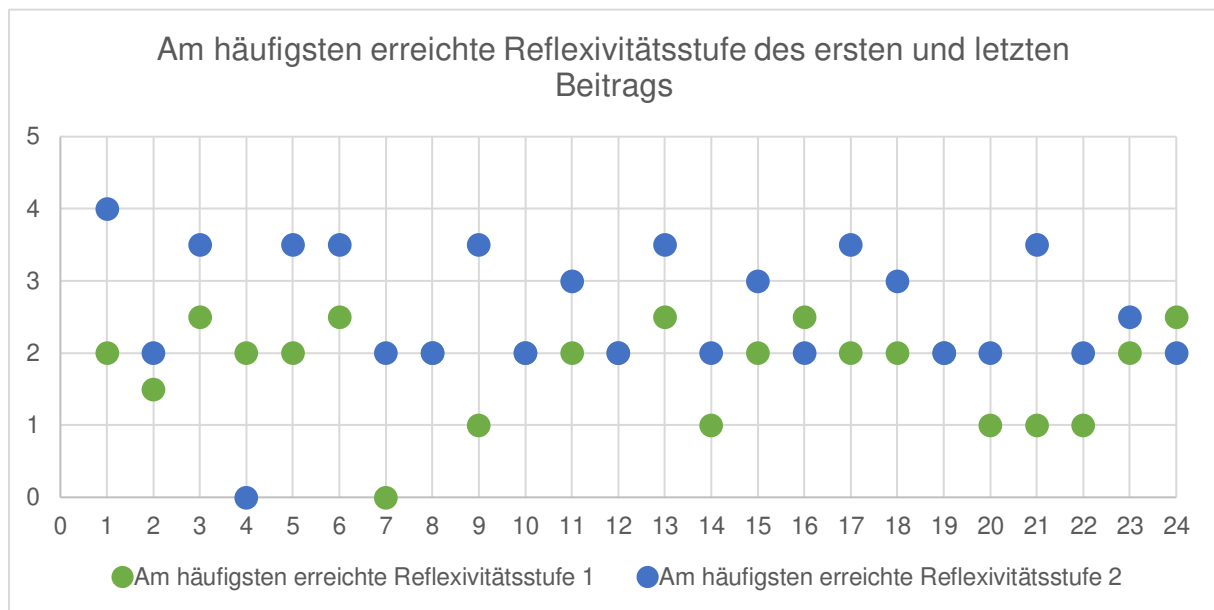
Abbildung 18 Am häufigsten erreichte Reflexionsstufe – erster und letzter Beitrag



Entwicklung der am häufigsten erreichten Reflexivitätsstufe

Mit Ausnahme von P7 erreichen alle Testpersonen mindestens Stufe 1 als am häufigsten erreichte Reflexivitätsstufe im ersten Blogbeitrag (vgl. Abbildung 19). Die fünf Teilnehmenden P3, P6, P13, P16 und P24 haben Stufe 2,5 als am häufigsten erreichte Reflexivitätsstufe. Die übrigen liegen zwischen den Stufen 1 und 2. Im letzten Blogbeitrag liegen die am häufigsten erreichten Reflexivitätsstufen bei den meisten Probandinnen und Probanden über den Stufen, die sie im ersten Beitrag zeigen. Bei P4, P16 und P24 liegen sie unter dem ersten Wert mit der Stufe 0 und einer Differenz von $-2,0$ bei P4 und der Stufe 2 und einer Differenz von $-0,5$ bei P16 und P24. P1 hat beim letzten Beitrag die höchste am häufigsten erreichte Reflexivitätsstufe von Stufe 4, die übrigen Testpersonen zeigen Stufenwerte zwischen 3,5 und 2 (Ausnahme P4) mit Differenzen zwischen 0,5 und 2,5. Keine Veränderung der am häufigsten erreichten Reflexivitätsstufe zwischen dem ersten und letzten Beitrag ist bei P8, P12 und P19 feststellbar.

Abbildung 19 Am häufigsten erreichte Reflexivitätsstufe – erster und letzter Beitrag



Zusammenfassend bleibt festzuhalten, dass überwiegend steigende Werte der Kriterien der Reflektierfähigkeiten vom ersten zum letzten Beitrag vorhanden sind. Die Intervention scheint folglich die Reflektierfähigkeiten positiv zu beeinflussen. Um diese Hypothese zu überprüfen, werden nachfolgend die Ergebnisse der Experimentalgruppe mit Ergebnissen einer Kontrollgruppe, die keine Intervention erfährt, in einer zusätzlichen quantitativen Analyse näher untersucht (vgl. Kapitel 3.5.3).

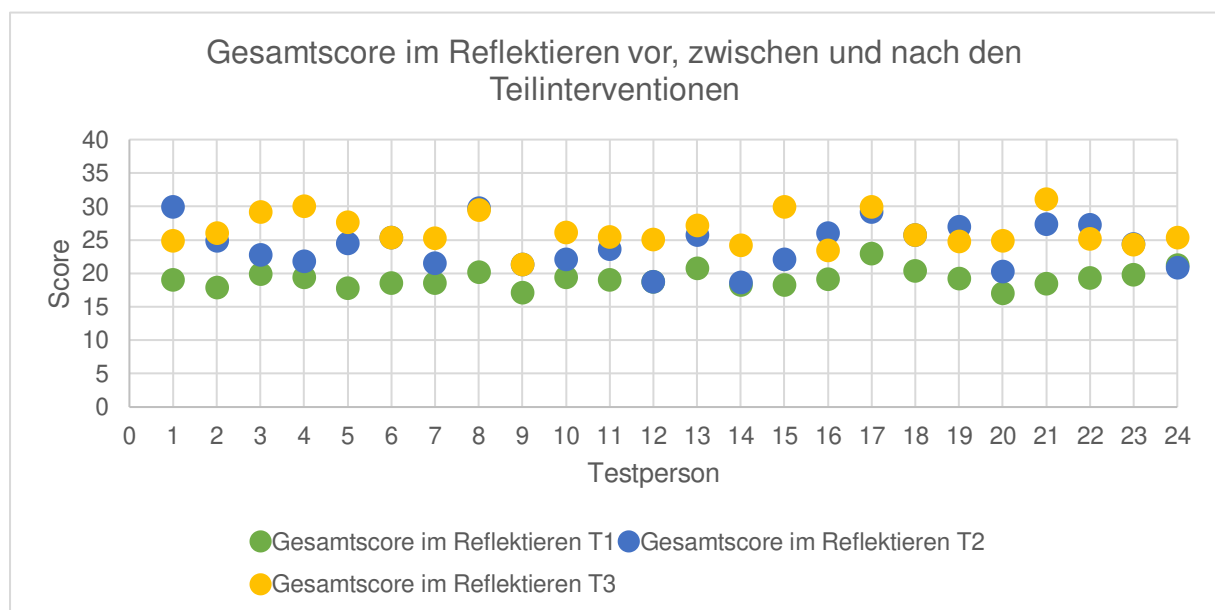
(4) Entwicklung der allgemeinen Reflektierfähigkeiten über beide Teilinterventionen hinweg

In diesem Kapitel werden die Daten zu den allgemeinen Reflektierfähigkeiten vor, zwischen und nach den zwei Teilinterventionen herangezogen und die Testpersonen hinsichtlich deren Entwicklungen in Gruppen eingeordnet. Die Daten zu den drei Messzeitpunkten bestehen aus den Durchschnittswerten der einzelnen Kriterien mehrerer Blogbeiträge, die in der jeweiligen Phase verfasst werden. Zunächst wird auf die drei Scores eingegangen und anschließend auf die maximal und am häufigsten erreichten Stufen. Im letzten Schritt werden die Entwicklungen der einzelnen Kriterien für die Reflektierfähigkeiten durch die erste Teilintervention als Untersuchungsgegenstand herangezogen, um die Schülerinnen und Schüler, die Gibbs Reflexivitätszyklus als erste Reflektierhilfe erhalten die Testpersonen, die die Hütemethode als erste Hilfe bekommen, gegenüberzustellen und sie hinsichtlich der gezeigten Entwicklungen in Gruppen einzuteilen.

Entwicklung des Gesamtscores im Reflektieren

Ähnlich wie beim Gesamtscore im Reflektieren des ersten Blogbeitrags sind auch die durchschnittlichen Werte des Gesamtscores im Reflektieren der zwei Beiträge vor der ersten Teilintervention, die zuhause verfasst werden. Den niedrigsten Ausgangswert hat P20 mit 17,0 und den höchsten P17 mit 22,9. Die Werte variieren folglich zwischen einem Wert von knapp 5 bei den Beiträgen vor der ersten Teilintervention (vgl. Abbildung 20). Nach der ersten Teilintervention, die für eine Hälfte der Testpersonen Gibbs Reflexivitätszyklus entspricht und für die andere Hälfte die Hütemethode ist, variieren die erreichten Werte des Gesamtscores zwischen 18,6 bei P14 und 29,9 bei P1. Die größte Veränderung des Gesamtscores im Reflektieren von Textblock 1 auf 2 und folglich nach der ersten Teilintervention zeigt P1 mit einer positiven Differenz von 10,9. P12 verändert sich von Textblock 1 auf 2 im Bereich des Gesamtscores nicht und bleibt weiterhin auf einem Wert von 18,7. Es muss jedoch auch angemerkt werden, dass diese Person am Tag der ersten Teilintervention krank war und deshalb keine erste Teilintervention erhalten hat. Eine Testperson (P24) zeigt sogar eine leicht negative Entwicklung von einem Wert von 21,2 auf 20,8 mit einer Differenz von - 0,4. Von Textblock 2 auf 3 und folglich nach der zweiten Teilintervention variieren die Differenzen des Gesamtscores im Reflektieren stark. Die positivste Veränderung hat P4 mit 8,3, die negativste Veränderung P1 mit - 5,1 und die geringste Veränderung haben mit einer Differenz von 0,0 P6, P9 und P19. Insgesamt sechs Testpersonen zeigen eine Verschlechterung des Gesamtscores nach der zweiten Teilintervention im Vergleich zur vorherigen Messung und 15 eine Verbesserung. Den höchsten Wert bei Textblock 3 hat P21 mit 31,0 und den geringsten Wert P9 mit 21,3.

Abbildung 20 Gesamtscore im Reflektieren - vor, zwischen und nach Teilinterventionen



Zusammenfassend kann festgehalten werden, dass alle Testpersonen mit Ausnahme dreier (P12, P14 und P24) deutliche positive Veränderung nach der ersten Teilintervention zeigen. Nach der zweiten Teilintervention fallen die Veränderungen unterschiedlich aus: Während manche Schülerinnen und Schüler am Ende den höchsten Gesamtscore im Reflektieren erreichen, fällt bei manchen Testpersonen der Gesamtscore im Reflektieren unter einen Wert von dem, der in Textblock 2 erreicht wird und bei wenigen bleibt der Wert ungefähr gleich. Folgende fünf Gruppen lassen sich für die Entwicklung des Gesamtscores im Reflektieren über die Intervention hinweg identifizieren: Teilnehmende, die bei Textblock 1 den geringsten Wert zeigen, nach der ersten Teilintervention einen höheren Wert und nach der zweiten Teilintervention einen noch höheren Wert. In dieser Gruppe befinden sich die Testpersonen P2, P3, P4, P5, P7, P10, P11, P13, P14, P15, P17, P20 und P21 (vgl. Tabelle 19). Weiterhin gibt es die Gruppe derjenigen, die bei Textblock 1 ebenfalls den geringsten Wert zeigen, nach der ersten Teilintervention den höchsten Wert und nach der zweiten Teilintervention einen Wert zwischen dem Ergebnis von Textblock 1 und 2. Hierzu gehören die Testpersonen P1, P8, P16, P19, P22 und P23. Zur dritten Gruppe zählen diejenigen, die sich zunächst verbessern und anschließend auf dem gleichen Niveau bleiben (P6, P9, P18). Gruppe 4 und 5 bilden diejenigen, die sich zunächst verschlechtern bzw. gleichbleiben und anschließend eine Verbesserung zeigen (P24, P12).

Tabelle 19 Übersicht Gruppeneinteilung in Hinblick auf die Entwicklung des Gesamtscores im Reflektieren

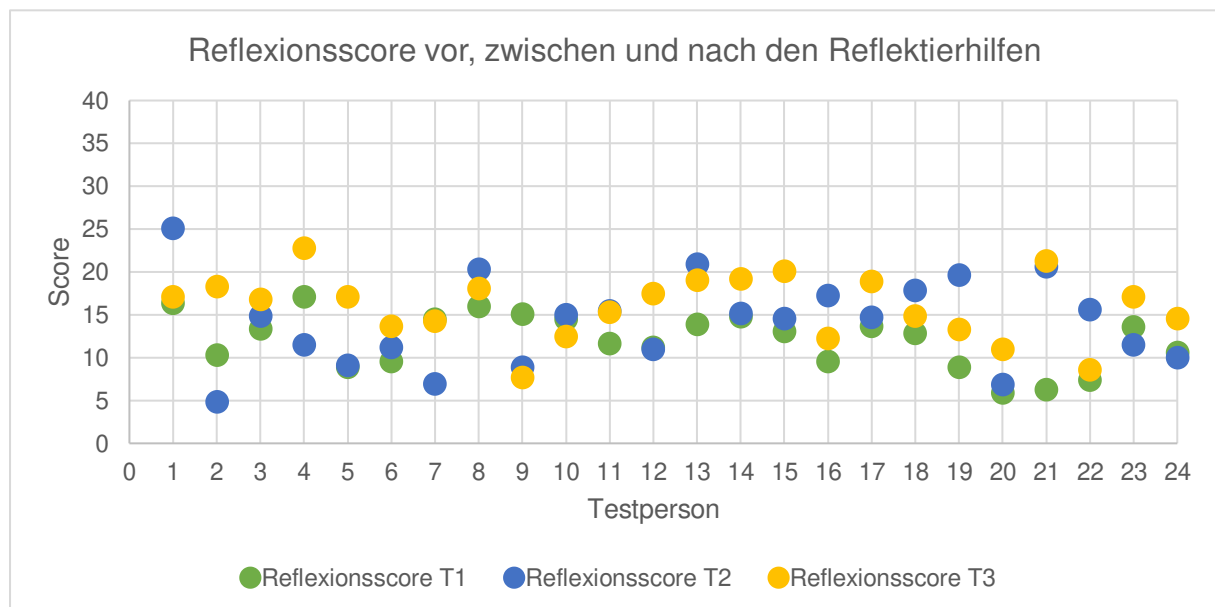
Gruppe 1 (Verbesserung T1 auf T2 und Verbesserung T2 auf T3)	<u>P2</u> , P3, <u>P4</u> , <u>P5</u> , <u>P7</u> , P10, P11, P13, <u>P14</u> , <u>P15</u> , <u>P17</u> , <u>P20</u> , P21
Gruppe 2 (Verbesserung T1 auf T2 und Verschlechterung T2 auf T3)	P1, P8, P16, P19, P22, <u>P23</u>
Gruppe 3 (Verbesserung T1 und T2 und gleichbleibend T2 auf T3)	<u>P6</u> , P9, P18
Gruppe 4 (Verschlechterung T1 auf T2 und Verbesserung T2 auf T3)	<u>P24</u>
Gruppe 5 (Gleichbleibend T1 auf T2 und Verbesserung T2 auf T3)	P12 ³¹

³¹ P12 war bei Erhalt der ersten Reflektierhilfe nicht anwesend, deshalb wird er durchgestrichen angezeigt.

Entwicklung des Reflexionsscores

Die Entwicklung des durchschnittlichen Reflexionsscores von den ersten beiden Beiträgen vor der ersten Teilintervention über die drei Beiträge danach und den drei Beiträgen nach der zweiten Teilintervention der Testpersonen ist stark unterschiedlich (vgl. Abbildung 21). Die durchschnittlichen Reflexionsscorewerte von Textblock 1 liegen zwischen 5,9 bei P20 und 1,1 bei P4. Nach der ersten Teilintervention liegen die Werte zwischen 25,1 bei P1 und 4,9 bei P2. 17 Probandinnen und Probanden haben im zweiten Textblock einen höheren Reflexionsscore als zuvor, die übrigen sieben Teilnehmenden zeigen nach der ersten Teilintervention niedrigere Werte als zuvor (P2, P4, P7, P9, P12, P23, P24). Die größte Veränderung des Reflexionsscore hat P21 mit einer Differenz von 14,3. P5 verändert seinen Wert mit einer Differenz von 0,2 nur leicht und P7 hat die größte negative Entwicklung des Reflexionsscores von Textblock 1 auf 2 mit einer Differenz von - 7,5. Nach der zweiten Teilintervention variiert der Reflexionsscore zwischen Werten von 7,7 bei P9 und 22,8 bei P4. Die größte positive Veränderung besitzt dabei P2 mit einer Differenz von 13,4 und die größte negative Veränderung P1 mit einer Differenz von - 8,0. P11 verändert seinen Reflexionsscore mit einer Differenz von - 0,2 am geringsten von allen Testpersonen von Textblock 2 auf 3. Zehn der 24 Personen zeigen nach der zweiten Teilintervention einen niedrigeren Reflexionsscore als nach der ersten Teilintervention (P1, P8, P9, P10, P11, P13, P16, P18, P19, P22), die übrigen Testpersonen weisen Zunahmen des Reflexionsscores auf.

Abbildung 21 Reflexionsscore - vor, zwischen und nach Teilinterventionen



Insgesamt lassen sich vier Gruppen von Reflexionsscore-Entwicklungen erkennen: Diejenigen, die nach beiden Teilinterventionen Verbesserungen bzw. Verschlechterungen im Reflexionsscore aufweisen, diejenigen, die nach der ersten Teilintervention eine Verbesserung und

nach der zweiten Teilintervention eine Verschlechterung des Reflexionsscores haben und diejenigen, die nach der ersten Teilintervention eine Verschlechterung und nach der zweiten Teilintervention eine Verbesserung zeigen. Die folgende Tabelle zeigt die Einteilung der Testpersonen in diese Einordnung:

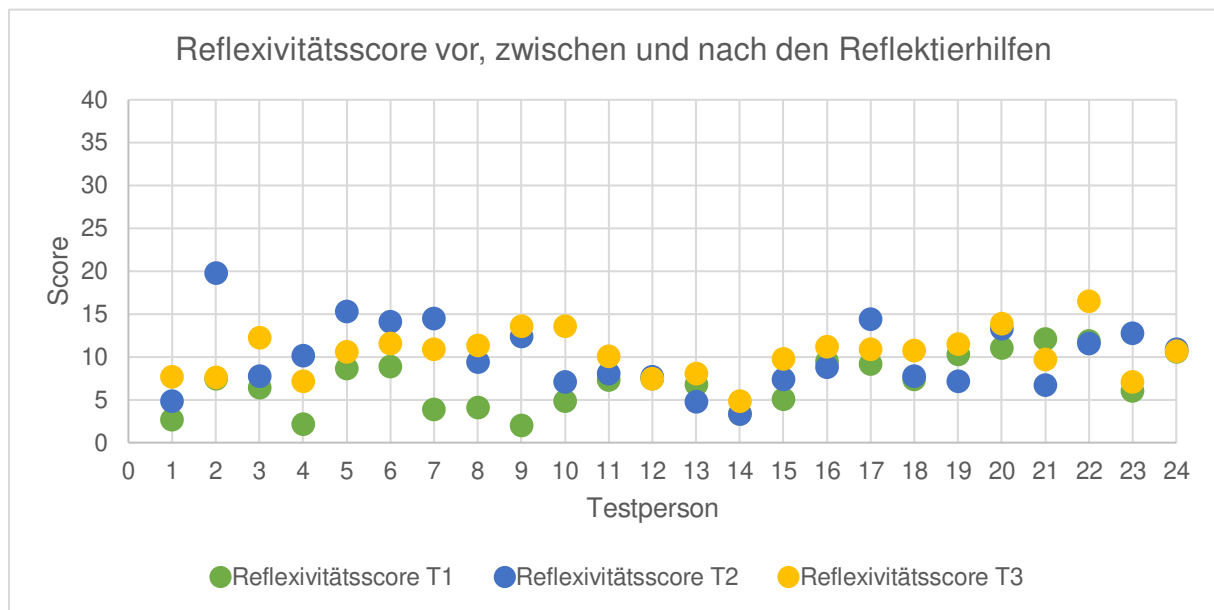
Tabelle 20 Übersicht Gruppeneinteilung in Hinblick auf die Entwicklung des Reflexionsscores

Gruppe 1 (Verbesserung T1 auf T2 und Verbesserung T2 auf T3)	P3, <u>P5</u> , <u>P6</u> , <u>P14</u> , <u>P15</u> , <u>P17</u> , <u>P20</u> , P21
Gruppe 2 (Verbesserung T1 auf T2 und Verschlechterung T2 auf T3)	P1, P8, P10, P11, P13, P16, P18, P19, P22
Gruppe 3 (Verschlechterung T1 auf T2 und Verbesserung T2 auf T3)	<u>P2</u> , <u>P4</u> , <u>P7</u> , P12 , <u>P23</u> , <u>P24</u>
Gruppe 4 (Verschlechterung T1 auf T2 und Verschlechterung T2 auf T3)	P9

Entwicklung des Reflexivitätsscores

Die Werte des Reflexivitätsscores im ersten Textblock sind ähnlich zu denen des ersten Beitrags. Sie variieren zwischen Werten von 12,1 bei P21 und 2,0 bei P9 (vgl. Abbildung 22). Im zweiten Textblock und folglich nach der ersten Teilintervention schwankt der Reflexivitätsscore um Werte zwischen 3,4 bei P14 und 19,8 bei P2. P2 zeigt nicht nur den höchsten Reflexivitätsscore zwischen beiden Teilinterventionen, sondern hat auch die größte Veränderung von Textblock 1 auf 2 mit einer Differenz von 12,3. Der Reflexivitätsscore von P14 hat keine Veränderung von Textblock 1 auf 2 und bei fünf Testpersonen sind negative Entwicklungen feststellbar mit einer maximalen negativen Differenz von -5,4 bei P21. Die Werte des Reflexivitätsscores beim dritten Textblock liegen zwischen 4,9 bei P14 und 16,5 bei P22. P2 entwickelt sich von Textblock 2 auf 3 mit einer Differenz des Reflexivitätsscores von -12,2 am negativsten, die anderen negativen Entwicklungen liegen zwischen einer Differenz von -5,7 und -0,2. P10 hat mit einer Differenz von 6,5 die größte positive Entwicklung des Reflexivitätsscores nach der zweiten Teilintervention. 15 Testpersonen zeigen von Textblock 2 auf 3 positive Veränderungen des Reflexivitätsscores und 14 negative Veränderungen.

Abbildung 22 Reflexivitätsscore - vor, zwischen und nach Teilinterventionen



Hinsichtlich des Reflexivitätsscores und dessen Entwicklung können die Probandinnen und Probanden – analog zum Reflexionsscore – in vier Gruppen eingeteilt werden (vgl. Tabelle 21). Gruppe 1 umfasst diejenigen, die nach beiden Teilinterventionen positive Veränderungen zeigen, Gruppe 2 diejenigen, die nach der ersten Teilintervention positive Veränderungen und nach der zweiten Teilintervention negative Entwicklungen haben, Gruppe 3 diejenigen, die nach der ersten Teilintervention negative Veränderungen und nach der zweiten Teilintervention positive Entwicklungen haben und Gruppe 4 diejenigen, die nach der ersten Teilintervention einen gleichen Reflexivitätsscore haben und diesen nach der zweiten Teilintervention verbessern.

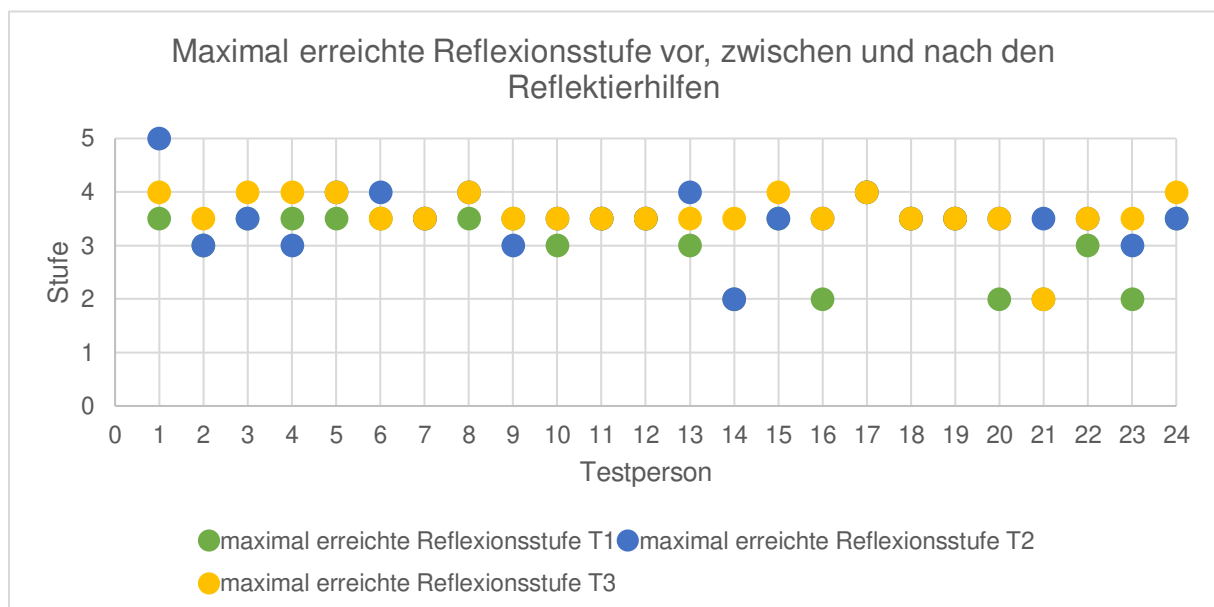
Tabelle 21 Übersicht Gruppeneinteilung in Hinblick auf die Entwicklung des Reflexivitätsscores

Gruppe 1 (Verbesserung T1 auf T2 und Verbesserung T2 auf T3)	P1, P3, P8, P9, P10, P11, <u>P15</u> , P18, <u>P20</u>
Gruppe 2 (Verbesserung T1 auf T2 und Verschlechterung T2 auf T3)	<u>P2</u> , <u>P4</u> , <u>P5</u> , <u>P6</u> , <u>P7</u> , <u>P12</u> , <u>P17</u> , <u>P23</u> , <u>P24</u>
Gruppe 3 (Verschlechterung T1 auf T2 und Verbesserung T2 auf T3)	P13, P16, P19, P21, P22
Gruppe 4 (Gleichbleibend T1 auf T2 und Verbesserung T2 auf T3)	<u>P14</u>

Entwicklung der maximal erreichten Reflexionsstufe

Die durchschnittlichen Werte der maximal erreichten Reflexionsstufe der Beiträge vor der ersten Teilintervention liegen zwischen 2,0 und 4,0 bei P17, wobei die Stufe 4 nur einmal vorkommt (vgl. Abbildung 23). Die übrigen Testpersonen erreichen bei Textblock 1 maximal Stufe 3,5. Im zweiten Textblock, der sich aus den durchschnittlichen Werten der drei Blogbeiträge zwischen den beiden Teilinterventionen zusammensetzt, weist maximal erreichte Reflexionsstufen zwischen 3 und 5 bei P1 auf. Die größte Stufenveränderung zeigen dabei P1, P16, P20 und P21 mit einer Stufendifferenz von 1,5. Zwei Teilnehmende (P4 und P9) erreichen in Textblock 2 geringere maximale Reflexionsstufen. Die übrigen Testpersonen zeigen keine Veränderung oder steigern die maximal erreichte Reflexionsstufe. Bei Textblock 3 liegen die maximal erreichten Reflexionsstufen zwischen 2 und 4, wobei Stufe 2 nur einmalig bei P21 vorkommt. Stufe 3 als maximal erreichte Reflexionsstufe hat keiner der Probandinnen und Probanden nach den beiden Teilinterventionen. Die größte positive Differenz von Textblock 2 auf 3 zeigt dabei P14 mit 1,5 und die größte negative Differenz P21 mit - 1,5. Der überwiegende Anteil der Testpersonen erreicht bei Textblock 3 die gleiche maximal erreichte Reflexionsstufe wie beim vorherigen Textblock oder eine positive Veränderung.

Abbildung 23 Maximal erreichte Reflexionsstufe - vor, zwischen und nach Teilinterventionen



Zusammenfassen lassen sich die Entwicklungen der maximal erreichten Reflexionsstufe der Testpersonen vor, zwischen und nach den Teilinterventionen in sechs Gruppen (vgl. Tabelle 22). Gruppe 1 besteht aus denjenigen, die zweimal eine Verbesserung zeigen, Gruppe 2 aus denen, die zunächst eine Verbesserung zeigen und dann einen konstanten Wert behalten, Gruppe 3 aus denjenigen, die zunächst eine Verbesserung zeigen und anschließend eine Verschlechterung. Gruppe 4 setzt sich aus den Testpersonen zusammen, die nach der ersten

Teilintervention die gleiche maximal erreichte Reflexionsstufe haben und nach der zweiten Teilintervention eine Verbesserung zeigen, Gruppe 5 aus denen, die über alle Zeiträume hinweg einen konstanten Wert haben und Gruppe 6 aus denjenigen, die zunächst eine Verschlechterung und im Anschluss eine Verbesserung in der maximal erreichten Reflexionsstufe aufweisen

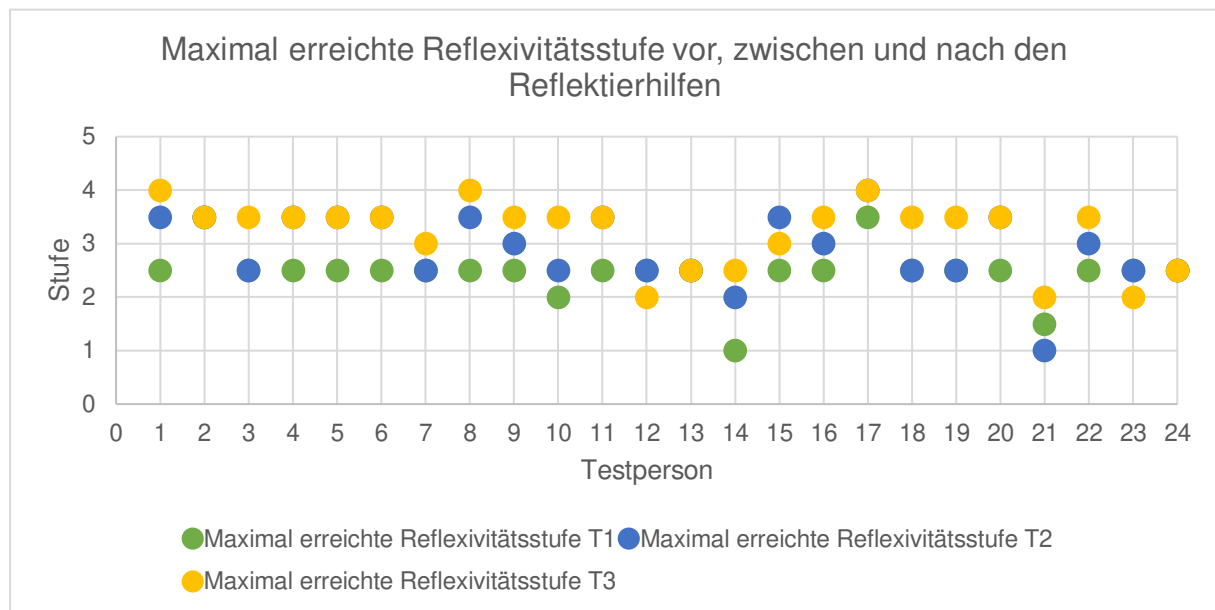
Tabelle 22 Übersicht Gruppeneinteilung in Hinblick auf die Entwicklung der maximal erreichten Reflexionsstufe

Gruppe 1 (Verbesserung T1 auf T2 und Verbesserung T2 auf T3)	<u>P23</u>
Gruppe 2 (Verbesserung T1 auf T2 und gleichbleibend T2 auf T3)	<u>P5</u> , P8, P10, P16, <u>P20</u> , P22
Gruppe 3 (Verbesserung T1 auf T2 und Verschlechterung T2 auf T3)	P1, <u>P6</u> , P13, P21
Gruppe 4 (Gleichbleibend T1 auf T2 und Verbesserung T2 auf T3)	<u>P2</u> , P3, <u>P14</u> , <u>P15</u> , <u>P24</u>
Gruppe 5 (Gleichbleibend T1 auf T2 und gleichbleibend T2 auf T3)	<u>P7</u> , P11, P12 , <u>P17</u> , P18, P19
Gruppe 6 (Verschlechterung T1 auf T2 und Verbesserung T2 auf T3)	<u>P4</u> , P9

Entwicklung der maximal erreichten Reflexivitätsstufe

Die maximal erreichte Reflexivitätsstufe der Testpersonen variiert bei Textblock 1 zwischen den Stufen 1 bei P14 und 3,5 bei P2 und P17, wobei die niedrigste Stufe nur einmal vorkommt, gefolgt von einmalig der Stufe 1,5 bei P21 und Stufe 2 bei P10 (vgl. Abbildung 24). Die übrigen Schülerinnen und Schüler haben alle als maximal erreichte Reflexivitätsstufe die Stufe 2,5 bei Textblock 1. Im zweiten Textblock gibt es Stufen zwischen 1 bei P21 und 4 bei P17. Die Stufe 1 gibt es nur einmal, ebenfalls Stufe 2 (P14) und Stufe 4. Die Stufen 3 und 3,5 werden am häufigsten als maximal erreichte Reflexivitätsstufen erreicht. Die größte Veränderung zwischen Textblock 1 und 2 haben mehrere Testpersonen mit einer Differenz von 1,0. P21 zeigt eine negative Entwicklung um den Wert - 0,5, die anderen behalten die gleiche Stufe oder steigern sich leicht. Im dritten Textblock sind die Stufen 2 bis 4 vertreten. Die am häufigsten vorkommende maximal erreichte Reflexivitätsstufe ist dabei Stufe 3,5. P12, P15 und P23 verschlechtern sich um eine halbe Stufe nach der zweiten Teilintervention. Die übrigen Testpersonen behalten das entsprechende Niveau oder steigern sich leicht.

Abbildung 24 Maximal erreichte Reflexivitätsstufe - vor, zwischen und nach Teilinterventionen



Analog zur Gruppeneinteilung der maximal erreichten Reflexionsstufe lassen sich die Testpersonen hinsichtlich der maximal erreichten Reflexivitätsstufe in sieben Gruppen einteilen:

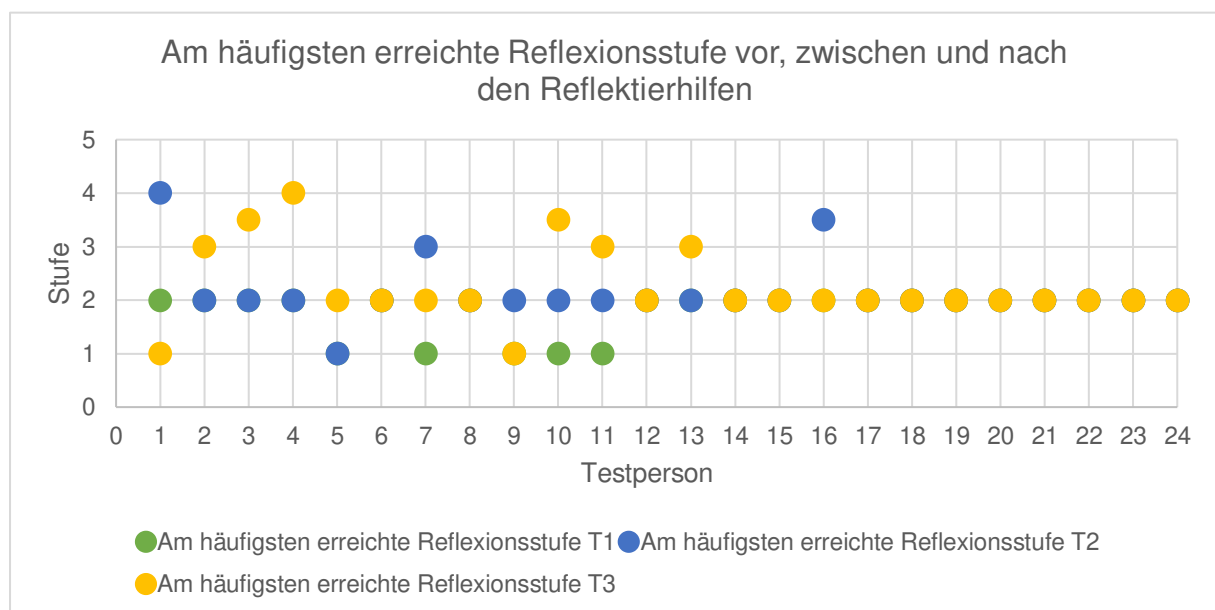
Tabelle 23 Übersicht Gruppeneinteilung in Hinblick auf die Entwicklung der maximal erreichten Reflexivitätsstufe

Gruppe 1 (Verbesserung T1 auf T2 und Verbesserung T2 auf T3)	P1, P8, P9, P10, <u>P14</u> , P16, P22
Gruppe 2 (Verbesserung T1 auf T2 und gleichbleibend T2 auf T3)	<u>P4</u> , <u>P5</u> , <u>P6</u> , P11, <u>P17</u> , <u>P20</u>
Gruppe 3 (Verbesserung T1 auf T2 und Verschlechterung T2 auf T3)	<u>P15</u>
Gruppe 4 (Gleichbleibend T1 auf T2 und Verbesserung T2 auf T3)	P3, <u>P7</u> , P18, P19
Gruppe 5 (Gleichbleibend T1 auf T2 und gleichbleibend T2 auf T3)	<u>P2</u> , P13, <u>P24</u>
Gruppe 6 (Gleichbleibend T1 auf T2 und Verschlechterung T2 auf T3)	<u>P12</u> , <u>P23</u>
Gruppe 7 (Verschlechterung T1 auf T2 und Verbesserung T2 auf T3)	P21

Entwicklung der am häufigsten erreichten Reflexionsstufe

Im ersten Textblock haben alle Testpersonen entweder Stufe 1 (P5, P7, P9, P10, P11) oder 2 als am häufigsten erreichte Reflexionsstufe (vgl. Abbildung 25). Die am häufigsten erreichte Reflexionsstufe im zweiten Textblock variiert zwischen Stufe 1 bei P5 und 4 bei P1. Die meisten Teilnehmenden haben dabei Stufe 2, P7 hat Stufe 3 und P16 Stufe 3,5. 21 Testpersonen haben bei Textblock 1 und 2 die gleichen am häufigsten erreichten Reflexionsstufen, P1 und P7 steigern sich um zwei Stufen und P16 um eineinhalb Stufen. Im dritten Textblock kommen ebenfalls die Stufe 1 bis 4 vor. P1, P5 und P9 haben dabei die niedrigsten Stufen vom Wert 1,0 und P4 die am höchsten erreichte Stufe der am häufigsten erreichten Reflexionsstufe vom Wert 4,0. P1, P7 und P9 haben im letzten Textblock einen niedrigeren Wert der am häufigsten erreichten Reflexionsstufe mit Differenzen von einmal den Wert 3,0 und zweimal den Wert 1,0. Die übrigen Testpersonen zeigen entweder die gleichen Stufen als zuvor oder verbessern sich um mindestens eine Stufe.

Abbildung 25 Am häufigsten erreichte Reflexionsstufe - vor, zwischen und nach Teilinterventionen



Die Probandinnen und Probanden können in Hinblick auf die am häufigsten erreichte Reflexionsstufe vor, zwischen und nach den Teilinterventionen in vier Gruppen eingeteilt werden: Gruppe 1 umfasst diejenigen, die nach beiden Teilinterventionen Verbesserungen zeigen, Gruppe 2 diejenigen, die zunächst eine Verbesserung und anschließend eine Verbesserung aufweisen, Gruppe 3 die Personen, die zunächst gleichbleibend sind und sich danach verbessern und Gruppe 4, diejenigen Testpersonen, die stets die gleiche Stufe besitzen (vgl. Tabelle 24).

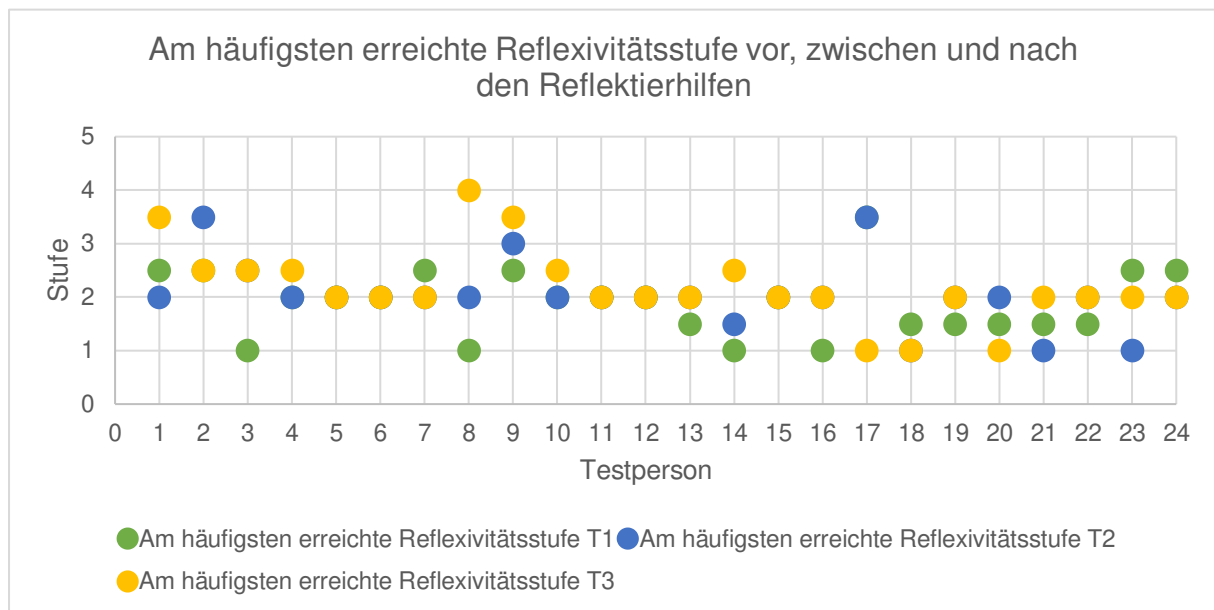
Tabelle 24 Übersicht Gruppeneinteilung in Hinblick auf Entwicklung der am häufigsten erreichten Reflexionsstufe

Gruppe 1 (Verbesserung T1 auf T2 und Verbesserung T2 auf T3)	P10, P11
Gruppe 2 (Verbesserung T1 auf T2 und Verschlechterung T2 auf T3)	P1, <u>P7</u> , P9, P16
Gruppe 3 (Gleichbleibend T1 auf T2 und Verbesserung T2 auf T3)	<u>P2</u> , P3, <u>P4</u> , <u>P5</u> , P13
Gruppe 4 (Gleichbleibend T1 auf T2 und gleichbleibend T2 auf T3)	P6, P8, <u>P12</u> , <u>P14</u> , <u>P15</u> , <u>P17</u> , P18, P19, <u>P20</u> , P21, P22, <u>P23</u> , <u>P24</u>

Entwicklung der am häufigsten erreichten Reflexivitätsstufe

Die am häufigsten erreichten Reflexivitätsstufen von Textblock 1 liegen zwischen den Stufen 1 und 3,5, wobei die höchste Stufe nur von P17 erreicht wird (vgl. Abbildung 26). P3, P8, P14 und P16 haben mit Stufe 1 die geringste am häufigsten erreichte Reflexivitätsstufe. Im zweiten Textblock variieren die Stufen zwischen 1 bei P21 und P23 und 3,5 bei P17. Bei P1, P21 und P23 sind negative Entwicklungen der am häufigsten erreichten Reflexivitätsstufe von Textblock 1 auf 2 feststellbar mit Differenzen von zweimal 0,5 und einmal 1,5. Acht Testpersonen haben bei beiden Textblöcken die gleiche Stufe, die übrigen verbessern sich hinsichtlich der am häufigsten erreichten Reflexivitätsstufe. Der dritte Textblock besteht aus Stufenwerten von 1 bei P17, P18 und P20 bis 4 bei P8. Dreizehn Testpersonen bleiben auf dem gleichen Stufenniveau wie bei Textblock 2, fünf steigern sich und drei verschlechtern sich hinsichtlich der am häufigsten erreichten Reflexivitätsstufe. Die größte Zunahme von Textblock 2 auf 3 hat P8 mit einer Differenz von 2 Stufen, die größte Abnahme P17 mit einer Differenz von 2,5 Stufen.

Abbildung 26 Am häufigsten erreichte Reflexivitätsstufe - vor, zwischen und nach Teilinterventionen



Hinsichtlich der am häufigsten erreichten Reflexivitätsstufe lassen sich insgesamt acht Testpersonengruppen identifizieren. Dabei gibt es Kombinationen aus Verbesserungen, Verschlechterungen und Gleichbleibungen von Textblock 1 auf 2 mit Verbesserungen, Verschlechterungen und Gleichbleibungen von Textblock 2 auf 3 mit Ausnahme der Kombination der doppelten Verschlechterung. Die Einteilung sieht wie folgt aus:

Tabelle 25 Übersicht Gruppeneinteilung in Hinblick auf die Entwicklung der am häufigsten erreichten Reflexivitätsstufe

Gruppe 1 (Verbesserung T1 auf T2 und Verbesserung T2 auf T3)	P8, P9, <u>P14</u>
Gruppe 2 (Verbesserung T1 auf T2 und Verschlechterung T2 auf T3)	<u>P2</u> , <u>P20</u>
Gruppe 3 (Verbesserung T1 auf T2 und gleichbleibend T2 auf T3)	P3, P13, P16, P19, P22
Gruppe 4 (Gleichbleibend T1 auf T2 und Verbesserung T2 auf T3)	<u>P4</u> , P10
Gruppe 5 (Gleichbleibend T1 auf T2 und gleichbleibend T2 auf T3)	<u>P5</u> , <u>P6</u> , P11, <u>P12</u> , <u>P15</u>

Gruppe 6 (Gleichbleibend T1 auf T2 und Verschlechterung T2 auf T3)	<u>P17</u>
Gruppe 7 (Verschlechterung T1 auf T2 und Verbesserung T2 auf T3)	P1, P21, <u>P23</u>
Gruppe 8 (Verschlechterung T1 auf T2 und gleichbleibend T2 auf T3)	<u>P7</u> , P18, <u>P24</u>

(5) Übersicht der Teilnehmenden über Entwicklung der Reflektierfähigkeiten in Abhängigkeit der jeweils erhaltenen Reflektierhilfe

Interessant ist der Einfluss der Teilinterventionen auf die Reflexions- und Reflexivitätsfähigkeiten. Deshalb werden die Ergebnisse der Testpersonen, die Gibbs Reflexivitätszyklus als erste Teilintervention erhalten und derjenigen, die die Hütemethode als erste Teilintervention erhalten, getrennt betrachtet und die Entwicklung der beiden Teilreflektierfähigkeiten der Probandinnen und Probanden vom ersten zum zweiten Messzeitpunkt und vom zweiten auf den dritten Messzeitpunkt in Abhängigkeit der erhaltenen Reflektierhilfe genauer vorgestellt. P12 wird hierbei ausgenommen, da die Testperson bei der ersten Teilintervention nicht anwesend war. Zunächst werden die Jugendlichen hinsichtlich der Reflexivitätsfähigkeit (zusammengesetzt aus Reflexivitätsscore und maximal und am häufigsten erreichte Reflexivitätsstufen) und anschließend hinsichtlich der Reflexionsfähigkeit (zusammengesetzt aus Reflexionsscore und maximal und am häufigsten erreichte Reflexionsstufen) vor und nach der ersten Teilintervention und vor und nach der zweiten Teilintervention vorgestellt.

Hinsichtlich der Förderung der Reflexivitätsfähigkeit werden die erhobenen Daten Reflexivitätsscore, maximal erreichte Reflexivitätsstufe und am häufigsten erreichte Reflexivitätsstufe zusammengefasst. Dabei lassen sich drei Testpersonengruppen identifizieren: Diejenigen, die bei allen drei Werten Steigerungen nach der ersten Teilintervention zeigen, diejenigen, die Verbesserungen und gleichbleibende Werte besitzen und diejenigen, die auch negative Entwicklungen aufweisen (vgl. Tabelle 26).

Tabelle 26 Gruppeneinteilung hinsichtlich der Entwicklung der Reflexivitätsfähigkeit nach erster Teilintervention

Gruppeneinteilung hinsichtlich der Entwicklung der Reflexivitätsfähigkeit	Personen, die Gibbs Reflexivitätszyklus als erste Teilintervention erhalten	Personen, die die Hütemethode als erste Teilintervention erhalten
Gruppe 1 (nur positive Entwicklungen aller drei Werte)	P20	P8, P9
Gruppe 2 (positive und keine Entwicklungen)	P2, P4, P5, P6, P14, P15, P17	P3, P10, P11
Gruppe 3 (auch negative Entwicklungen)	P7, P23, P24	P1, P13, P16, P18, P19, P21, P22

Zur Untersuchung der Entwicklung der Reflexionsfähigkeiten werden die Ergebnisse des Reflexionsscores, der maximal erreichten Reflexionsstufe und der am häufigsten erreichten Reflexionsstufe herangezogen. Es lassen sich ebenfalls drei Gruppen an Testpersonen hinsichtlich der Entwicklung dieser drei Werte identifizieren (vgl. Tabelle 27). Vergleicht man die Einteilung der Probandinnen und Probanden, die Gibbs Reflexivitätszyklus als erste Teilintervention erhalten, mit denjenigen, die die Hütemethode zunächst erhalten, stellt man fest, dass erstere bessere Ergebnisse hinsichtlich der Reflexivitätsfähigkeit zeigen und letztere hinsichtlich der Reflexionsfähigkeit.

Tabelle 27 Gruppeneinteilung hinsichtlich der Entwicklung der Reflexionsfähigkeit nach erster Teilintervention

Gruppeneinteilung hinsichtlich der Entwicklung der Reflexionsfähigkeit	Personen, die Gibbs Reflexivitätszyklus als erste Teilintervention erhalten	Personen, die die Hütemethode als erste Teilintervention erhalten
Gruppe 1 (nur positive Entwicklungen aller drei Werte)	/	P1, P10, P16
Gruppe 2 (positive und keine Entwicklungen)	P5, P6, P14, P15, P17, P20	P3, P8, P11, P13, P18, P19, P21, P22
Gruppe 3 (auch negative Entwicklungen)	P2, P4, P7, P23, P24	P9

Die zweite Teilintervention in Form von Gibbs Reflexivitätszyklus bewirkt eine Steigerung aller drei Reflexivitätskriterien bei fünf Testpersonen, die übrigen sechs zeigen neben positiven Entwicklungen auch mindestens einmal keine Veränderung (vgl. Tabelle 28). Es gibt keine Person, die Gibbs als zweite Teilintervention erhält und eine negative Entwicklung hinsichtlich der Reflexivitätsfähigkeit zeigt. Bei denjenigen, die als zweite Teilintervention die Hütemethode

bekommen, haben zehn von 12 Testpersonen mindestens eine negative Veränderung bei einer der drei Reflexivitätskriterien. Eine Person hat nur positive Entwicklungen und eine weitere zeigt positive und negative Entwicklungen hinsichtlich der Reflexivitätsfähigkeit. Gibbs Reflexivitätszyklus wirkt sich folglich positiv auf die Reflexivitätsfähigkeit auf, während die Hütemethode überwiegend zu Verschlechterungen dieser Fähigkeit führt.

Tabelle 28 Gruppeneinteilung hinsichtlich der Entwicklung der Reflexivitätsfähigkeit nach zweiter Teilintervention

Gruppeneinteilung hinsichtlich der Entwicklung der Reflexivitätsfähigkeit	Personen, die Gibbs Reflexivitätszyklus als zweite Teilintervention erhalten	Personen, die die Hütemethode als zweite Teilintervention erhalten
Gruppe 1 (nur positive Entwicklungen aller drei Werte)	P1, P8, P9, P10, P21	P14
Gruppe 2 (positive und keine Entwicklungen)	P3, P11, P13, P16, P18, P19, P22	P24
Gruppe 3 (auch negative Entwicklungen)	/	P2, P4, P5, P6, P7, P15, P17, P20, P23

Die Reflexionsfähigkeit entwickelt sich hingegen gegensätzlich zur Reflexivitätsfähigkeit durch die zweite Teilintervention. Bei den Testpersonen, die Gibbs Reflexivitätszyklus als zweite Reflektierhilfe erhalten, gibt es mit Ausnahme einer Person, jeweils mindestens eine negative Entwicklung hinsichtlich der drei Reflexionskriterien Score, maximal erreichte Stufe und am häufigsten erreichte Stufe (vgl. Tabelle 29). Bei den Testpersonen, die zu diesem Zeitpunkt die Hütemethode erhalten, zeigen die meisten sowohl positive als auch keine Entwicklungen und jeweils zwei Personen weisen nur positive oder auch negative Entwicklungen auf.

Tabelle 29 Gruppeneinteilung hinsichtlich der Entwicklung der Reflexionsfähigkeit nach zweiter Teilintervention

Gruppeneinteilung hinsichtlich der Entwicklung der Reflexionsfähigkeit	Personen, die Gibbs Reflexivitätszyklus als zweite Teilintervention erhalten	Personen, die die Hütemethode als zweite Teilintervention erhalten
Gruppe 1 (nur positive Entwicklungen aller drei Werte)	P3	P2, P4
Gruppe 2 (positive und keine Entwicklungen)	/	P5, P14, P15, P17, P20, P23, P24
Gruppe 3 (auch negative Entwicklungen)	P1, P8, P9, P10, P11, P13, P16, P18, P19, P21, P22	P6, P7

Zusammengefasst lassen sich folglich je nach erhaltene erste Reflektiermethode unterschiedliche Entwicklungen der Testpersonen erkennen. Diese Beobachtung und die Erkenntnisse der Untersuchung der Schülerinnen und Schüler bezüglich ihrer Entwicklung der Reflektierfähigkeiten und Begriffsverständnisse vom ersten zum letzten Beitrag sollen nachfolgend mithilfe statistischer Auswertungsverfahren genauer untersucht werden, um konkretere Aussagen über die Wirkungen der Gesamtintervention und der Teilinterventionen treffen zu können, die im Rahmen der qualitativen Auswertung bereits erkenntlich wurden.

3.5.3 Ergebnisse der quantitativen Auswertung

Im Folgenden werden die Ergebnisse der Varianzanalysen zur Testung der aufgestellten Hypothesen vorgestellt. Dabei wird sowohl auf Signifikanzen als auch Effektstärken eingegangen. Anzumerken ist, dass sich die deskriptive Beschreibung im Rahmen dieser quantitativen Auswertung auf Übersichten von Mittelwerten, Minima, Maxima und Varianzen in tabellarischer Form reduziert. Grund dafür ist, dass bereits im Rahmen der qualitativen Auswertung ein Überblick über einzelne Werte gegeben wurde. Alle Tabellen für die deskriptive Beschreibung der Daten, die für die Hypothesentestung und -generierung von Bedeutung sind, befinden sich im Anhang (vgl. Anhang, S. A46). Ebenfalls werden die Voraussetzungen für Varianzanalysen bei jeder Hypothesentestung und -generierung geprüft. Bei Verletzung der Varianzhomogenität werden anstelle von ANOVAs³² Welch-Tests durchgeführt, gegenüber Verletzungen anderer Voraussetzungen sind Varianzanalysen sehr robust, sodass diesen keine weitere Beachtung erfährt.

3.5.3.1 Hypothese 1

Wissenschaftliche Hypothese 1:	Die Intervention bestehend aus zwei Teilinterventionen führt zu einer gesteigerten Reflektierfähigkeit.	
Statistisch:	H ₀ : Die Reflektierfähigkeiten der drei Faktorstufen Experimentalgruppe 1 und 2 und Kontrollgruppe verbessern sich vom Prä- zum Posttest nicht signifikant.	H ₁ : Die Reflektierfähigkeiten der Experimentalgruppe 1 und 2 verbessern sich vom Prä- zum Posttest signifikant, während sich die Reflektierfähigkeiten der Kontrollgruppe nicht signifikant verändern.

³² ANOVA ist die Kurzform von Varianzanalyse und leitet sich vom englischen Begriff ab.

Als erstes Ergebnis bleibt festzuhalten, dass es eine statistisch signifikante, lineare Interaktion zwischen dem Messzeitpunkt und den Zwischensubjektfaktorstufen bei allen Variablen mit Ausnahme der Variable Reflexivitätsscore mit $F(2, 36) = 2,76$, $p < ,076$, partielles $\eta^2 = ,13$ ³³ gibt. Die Signifikanzen liegen bei den Variablen Gesamtscore, Reflexionsscore und maximal erreichte Reflexivitätsstufe bei $p < ,001$. Die Variable maximal erreichte Reflexionsstufe hat eine signifikante Interaktion von $p = ,003$, die Variable am häufigsten erreichte Reflexivitätsstufe von $p = 0,008$ und die Variable am häufigsten erreichte Reflexionsstufe von $p = 0,010$. Da Interaktionseffekte vorliegen, werden nun die einfachen Haupteffekte des Zwischensubjektfaktors für die entsprechenden Variablen vorgestellt. Dafür werden die Ergebnisse der Welch-Analysen und des Games-Howell Tests als post-hoc Test genauer betrachtet, da die Varianzhomogenität als Voraussetzung für Varianzanalysen nicht bei allen Variablen gewährleistet ist. Keine der Variablen Gesamtscore im Reflektieren, Reflexionsscore, maximal erreichte Reflexionsstufe, maximal erreichte Reflexivitätsstufe und am häufigsten erreichte Reflexivitätsstufe haben einen signifikanten Unterschied zwischen den drei Faktorstufen zum ersten Messzeitpunkt. Folglich fallen die post-hoc Vergleiche weg, da keine signifikanten Unterschiede näher untersucht werden müssten. Lediglich für die Variable am häufigsten erreichte Reflexionsstufe, für die keine Welch ANOVA aufgrund der Tatsache, dass in mindestens einer Gruppe eine Varianz von Null auftritt, berechenbar ist, wird der Games-Howell Test betrachtet. Auch hier zeigen sich zwischen den Faktorstufen keine signifikanten Unterschiede. Die Signifikanz zwischen Faktorstufe 1 und 2 beträgt $p = ,180$, die zwischen Faktorstufe 2 und 3 $p = ,841$ und die zwischen Faktorstufe 1 und 3 $p = ,559$. Für die Variable Reflexivitätsscore wird der Haupteffekt der Gruppe betrachtet. Dieser beträgt $F(2, 36) = 3,94$ mit $p = ,028$ und partielles $\eta^2 = ,18$. Der Welch-Test zeigt, dass der Reflexivitätsscore des ersten Beitrags ebenfalls keinen signifikanten Unterschied zwischen den Faktorstufen der Gruppe aufweist mit $p = ,380$. Die Ausgangsbedingungen der drei Faktorstufen sind folglich alle miteinander vergleichbar und keine der Variablen des ersten Beitrags unterscheidet sich signifikant in Hinblick auf den Faktor Gruppe.

Alle Variablen zeigen hingegen signifikante Unterschiede zwischen den drei Faktorstufen zu Messzeitpunkt 2. Die höchsten Signifikanzen sind dabei bei den Variablen Gesamtscore im Reflektieren, maximal erreichte Reflexionsstufe und maximal erreichte Reflexivitätsstufe mit jeweils $p = ,000$ zu finden, gefolgt von den Variablen am häufigsten erreichte Reflexivitätsstufe mit $p = ,001$ und Reflexionsscore mit $p = ,002$. Die Signifikanz der Variable am häufigsten erreichte Reflexionsstufe beträgt $p = ,005$ und die der Variable Reflexivitätsscore $p = ,007$. Letzteres weist zwar keinen Interaktionseffekt auf, jedoch einen Haupteffekt ($F(2, 36) = 3,94$ mit $p = ,028$ und partielles $\eta^2 = ,18$).

³³ Im Rahmen der statistischen Auswertung wird die Null vorm Komma weggelassen.

Da bei allen abhängigen Variablen signifikante Unterschiede zwischen den Faktorstufen zu Messzeitpunkt 2 bestehen, werden nun die Ergebnisse des Games-Howell Tests als post-hoc Test vorgestellt. Alle Variablen zeigen zwischen den Faktorstufen 1 und 2 (beide Experimentalgruppen mit entweder Gibbs oder Hütemethode als erste Teilintervention) keine signifikanten Unterschiede ($p > ,050$). Zwischen den Faktorstufen 1 und 3 gibt es bei allen Variablen signifikante Unterschiede, zwischen den Faktorstufen 2 und 3 nur bei manchen Variablen. Die Variable Gesamtscore im Reflektieren unterscheidet sich sowohl zwischen Faktorstufe 1 und 3 (-9,48, 95%-CI[-11,86, -7,11]) als auch zwischen Faktorstufe 2 und 3 (-11,67, 95%-CI[-14,31, -9,03]) deutlich signifikant mit jeweils $p < ,001$. Auch die Variablen Reflexionsscore und maximal erreichte Reflexionsstufe unterscheiden sich sowohl zwischen Faktorstufe 1 und 3 (Variable 1: -6,32, 95%-CI[-11,15, -1,49]; Variable 2: -0,82, 95%-CI[-1,33, -,30]) als auch zwischen Faktorstufe 2 und 3 (Variable 1: -6,57, 95%-CI[-11,27, -1,88]; Variable 2: -,98, 95%-CI[-1,55, -,42]) signifikant mit $p < ,050$. Zwischen den Faktorstufen 1 und 3, d.h. zwischen den Testpersonen, die Gibbs Reflexivitätszyklus als erste Teilintervention erhalten und denjenigen, die keine Intervention erfahren, gibt es bei den Variablen Reflexivitätsscore, maximal erreichte Reflexivitätsstufe, am häufigsten erreichte Reflexionsstufe und am häufigsten erreichte Reflexivitätsstufe keine signifikanten Unterschiede mit $p > ,050$.

Zusammengefasst kann festgehalten werden, dass sich zum zweiten Messzeitpunkt alle Variablen zwischen Faktorstufe 2 und 3 signifikant unterscheiden und alle Variablen zwischen Faktorstufe 1 und 2 nicht signifikant unterscheiden. Die Variablen zwischen Faktorstufe 1 und 3 sind teilweise signifikant verschieden, teilweise auch nicht. Die Variablen beider Experimentalgruppen sind folglich ähnlich, sodass schlussgefolgert werden kann, dass die Intervention ähnliche Effekte auf die beiden Gruppen hat.

In Hinblick auf den Innersubjektfaktor Messzeitpunkt zeigt sich, dass die Gruppen 1 und 2 signifikante Effekte von Zeit auf die Variablen haben ($p < ,001$; partielles $\eta^2 = ,99$ bei Gruppe 1 und partielles $\eta^2 = ,98$ bei Gruppe 2). Gruppe 3 zeigt keinen allgemeinen Effekt von Zeit auf die Variablen mit $p = ,118$ und partielles $\eta^2 = ,61$. Auf die Variable Reflexivitätsscore, die keinen Interaktionseffekt Gruppe \times Zeit hat, ist ein signifikanter Haupteffekt von Zeit mit $F(1, 36) = 8,520$, $p = ,006$ und partielles $\eta^2 = ,19$ feststellbar.

Im Folgenden soll näher darauf eingegangen werden, welche Variablen genau Differenzen innerhalb der einzelnen Faktorstufen aufweisen. Bei Faktorstufe 1 (Gibbs als erste Teilintervention) gibt es auf alle Variablen mit Ausnahme der Variablen maximal erreichte Reflexivitätsstufe und am häufigsten erreichte Reflexivitätsstufe statistisch signifikante Effekte vom Messzeitpunkt. Die Variable Gesamtscore im Reflektieren zeigt einen signifikanten Effekt von $F(1, 11) = 311,56$, $p < ,001$, die Variable Reflexionsscore von $F(1, 11) = 40,99$, $p < ,001$, die Variable Reflexivitätsscore von $F(1, 11) = 7,33$, $p = ,020$, die Variable maximal erreichte Re-

flexionsstufe von $F(1, 11) = 15,82$, $p = ,002$ und die Variable am häufigsten erreichte Reflexionsstufe von $F(1, 11) = 9,16$, $p = ,012$. Auch bei Faktorstufe 2 (Hütemethode als erste Teilintervention) sind auf alle Variablen mit Ausnahme der Variable Reflexivitätsscore statistisch signifikante Effekte vom Messzeitpunkt feststellbar. Der signifikante Effekt der Variable Gesamtscore im Reflektieren beträgt $F(1, 11) = 133,66$, $p < ,001$, der der Variable Reflexionsscore $F(1, 11) = 20,64$, $p = ,001$, der der Variable maximal erreichte Reflexionsstufe $F(1, 11) = 12,79$, $p = ,004$, der der Variable maximal erreichte Reflexivitätsstufe $F(1, 11) = 98,79$, $p < ,001$, der der Variable am häufigsten erreichte Reflexionsstufe $F(1, 11) = 10,65$, $p = ,008$ und der der Variable am häufigsten erreichte Reflexivitätsstufe $F(1, 11) = 11,30$, $p = ,006$. Die Faktorstufe 3 (keine Intervention) weist nur einen signifikanten Effekt vom Messzeitpunkt auf die Variable Gesamtscore im Reflektieren auf mit $F(1, 14) = 13,37$, $p = ,003$. Alle anderen Variablen sind nicht von statistisch signifikanten Effekten betroffen.

Anhand der paarweisen Vergleiche innerhalb der Faktorstufen kann untersucht werden, in welche Richtung die Entwicklungen der Variablen gehen. Dabei stellt sich heraus, dass die Differenzen der sich signifikant unterscheidenden Variablen des ersten und letzten Beitrags der Faktorstufen 1 und 2 positiv gerichtet sind, d.h. dass die zweiten Werte signifikant größer sind als zuvor. Die Differenzen der Variablen der Kontrollgruppe (Faktorstufe 3) mit Ausnahme der Variable Gesamtscore im Reflektieren sind nicht signifikant, sodass hierfür paarweise Vergleiche nicht von Bedeutung sind. Die Variable Gesamtscore im Reflektieren der Faktorstufe 3 besitzt eine negativ gerichtete Differenz erkennbar an der mittleren Differenz von 1,29 (Wert von Messzeitpunkt 1 minus Wert von Messzeitpunkt 2).

Im Folgenden werden die Effektstärken nach d Cohen miteinander verglichen (vgl. Tabelle 30): Die größten Effekte der Intervention sind auf die Variable Gesamtscore im Reflektieren bei beiden Experimentalgruppen erkennbar mit $d = 4,061$ bei Experimentalgruppe 1 ($M_1 = 19,02$ und $M_2 = 28,13$) und $d = 4,772$ bei Experimentalgruppe 2 ($M_1 = 19,18$ und $M_2 = 30,31$). Experimentalgruppe 1 hat ebenfalls hohe Effekte bei den Variablen maximal erreichte Reflexionsstufe mit $d = 1,458$ ($M_1 = 2,96$ und $M_2 = 3,75$), Reflexionsscore mit $d = 1,442$ ($M_1 = 12,27$ und $M_2 = 18,88$) und am häufigsten erreichte Reflexionsstufe mit $d = 1,051$ ($M_1 = 1,75$ und $M_2 = 2,46$). Ein mittlerer Effekt ist auf die Variable Reflexivitätsscore mit $d = ,617$ ($M_1 = 6,77$ und $M_2 = 9,27$) vorhanden. Bei Experimentalgruppe 2 zeigen sich zusätzlich zur Variable Gesamtscore im Reflektieren hohe Effekte auf den Reflexionsscore mit $d = 1,792$ ($M_1 = 11,18$ und $M_2 = 19,13$), die maximal erreichte Reflexionsstufe mit $d = 1,745$ ($M_1 = 2,79$ und $M_2 = 3,92$), die am häufigsten erreichte Reflexivitätsstufe mit $d = 1,405$ ($M_1 = 1,88$ und $M_2 = 2,83$) und die am häufigsten erreichte Reflexionsstufe mit $d = 1,333$ ($M_1 = 2,00$ und $M_2 = 2,92$). Die maximal erreichte Reflexivitätsstufe hat einen mittleren Effekt von $d = ,733$ ($M_1 = 2,17$ und $M_2 = 3,46$). Bei der Kontrollgruppe, die keine Intervention erfährt, entwickelt sich der Gesamtscore im Reflektieren negativ mit $d = - ,629$ ($M_1 = 19,93$ und $M_2 = 18,64$).

Tabelle 30 Effektstärken auf die Variablen vom ersten zum letzten Beitrag aller drei Faktorstufen

Variable	Experimental- gruppe 1	Experimental- gruppe 2	Kontrollgruppe
Gesamtscore im Reflektieren	4,061	4,772	- ,629
Reflexionsscore	1,442	1,792	/
Reflexivitätsscore	,617	/	/
Maximal erreichte Reflexionsstufe	1,459	1,745	/
Maximal erreichte Reflexivitätsstufe	/	,733	/
Am häufigsten erreichte Reflexionsstufe	1,051	1,333	/
Am häufigsten erreichte Reflexivitätsstufe	/	1,405	/

Die erste Hypothese bestätigt sich mit der durchgeführten multivariaten zweifaktoriellen Varianzanalyse und die Nullhypothese wird abgelehnt. Während alle Gruppen zunächst ähnliche abhängige Variablenwerte zum ersten Messzeitpunkt haben, unterscheiden sich beim zweiten Messzeitpunkt vor allem die Messwerte der beiden Experimentalgruppen von denen der Kontrollgruppe deutlich. Bei genauerer Untersuchung der Differenzen der Variablen innerhalb der Gruppen stellt sich heraus, dass mit Ausnahme der Variablen am häufigsten erreichte und maximal erreichte Reflexivitätsstufe bei der Experimentalgruppe, die Gibbs Reflexivitätszyklus als erste Teilintervention erhält, signifikante Unterschiede feststellbar sind. Bei der Gruppe, die die Hütemethode als erste Teilintervention bekommt, unterscheiden sich alle Variablen außer der Variable Reflexivitätsscore des ersten Messzeitpunkts zum letzten Messzeitpunkt deutlich. Die Variablen der Kontrollgruppe ändern sich bis auf den Gesamtscore im Reflektieren (mit negativer Entwicklung) vom ersten zum zweiten Messzeitpunkt nicht signifikant.

3.5.3.2 Hypothese 2 und 3

Wissenschaftliche Hypothese 2:	Die Teilintervention in Form von Gibbs Reflexivitätszyklus führt zu einer Verbesserung der Reflexivitätsfähigkeit.	
Statistisch:	H ₀ : Die Reflexivitätsfähigkeiten aller Teilnehmenden verbessern sich nicht signifikant von Messzeitpunkt 1 auf 2 und von Messzeitpunkt 2 auf 3.	H ₁ : Die Reflexivitätsfähigkeiten von Messzeitpunkt 1 auf 2 und von Messzeitpunkt 2 auf 3 verbessern sich signifikant bei den Teilnehmenden, die zwischen den beiden Messzeitpunkten Gibbs Reflexivitätszyklus als Teilintervention erhalten, während die Reflexivitätsfähigkeiten der anderen Teilnehmenden keine signifikanten Unterschiede aufweisen.

Wissenschaftliche Hypothese 3:	Die Teilintervention in Form der Hütemethode führt zu einer Verbesserung der Reflexionsfähigkeit.	
Statistisch:	H ₀ : Die Reflexionsfähigkeiten aller Teilnehmenden verbessern sich nicht signifikant von Messzeitpunkt 1 auf 2 und von Messzeitpunkt 2 auf 3.	H ₁ : Die Reflexionsfähigkeiten von Messzeitpunkt 1 auf 2 und von Messzeitpunkt 2 auf 3 verbessern sich signifikant bei den Teilnehmenden, die zwischen den beiden Messzeitpunkten die Bonos Hütemethode als Teilintervention erhalten, während die Reflexionsfähigkeiten der anderen Teilnehmenden keine signifikanten Unterschiede aufweisen.

Für die meisten Variablen existiert eine statistisch signifikante Interaktion zwischen dem Messzeitpunkt und den Gruppen. Eine quadratisch statistisch signifikante Interaktion ist bei den Variablen Gesamtscore im Reflektieren ($F(1, 21) = 6,59, p = ,018$, partielles $\eta^2 = ,24$), Reflexionsscore ($F(1, 21) = 33,91, p < ,001$, partielles $\eta^2 = ,62$), Reflexivitätsscore ($F(1, 21) = 15,81, p = ,001$, partielles $\eta^2 = ,43$) und maximal erreichte Reflexivitätsstufe ($F(1, 21) = 8,82, p < ,007$, partielles $\eta^2 = ,30$) feststellbar. Bei der Variable am häufigsten erreichte Reflexivitätsstufe gibt es eine lineare statistisch signifikante Interaktion zwischen dem Messzeitpunkt und den beiden Gruppen mit $F(1, 21) = 7,19, p = ,014$ und dem partiellen $\eta^2 = ,26$. Die Variablen am häufigsten erreichte Reflexionsstufe und maximal erreichte Reflexionsstufe besitzen weder eine quadratische noch eine lineare statistisch signifikante Interaktion.

Im Folgenden werden die einfachen Haupteffekte des Zwischensubjektfaktors Gruppe durch einen Welch-Test für die Variablen mit Interaktionseffekten vorgestellt. Die Variablen Gesamtscore im Reflektieren, Reflexionsscore und Reflexivitätsscore unterscheiden sich nach der ersten Teilintervention signifikant ($p = ,036, p < ,001, p = ,010$), während sie vorher keine signifikanten Unterschiede aufweisen. Auch nach der zweiten Teilintervention sind die Ergebnisse wieder ohne signifikante Unterschiede. Die Variable maximal erreichte Reflexivitätsstufe mit $p_1 = ,442, p_2 = ,117$ und $p_3 = ,338$ zeigt zwischen beiden Faktorstufen zu keinem Messzeitpunkt signifikante Unterschiede, ebenso die Variablen maximal erreichte Reflexionsstufe und am häufigsten erreichte Reflexionsstufe, die sowieso keine Interaktionseffekte zeigen und auch keine Haupteffekte aufweisen ($F(1, 21) = ,041$ mit $p = ,842$ und $F(1, 21) = ,216$ mit $p = ,615$). Zum ersten Messzeitpunkt zeigt lediglich die Variable am häufigsten erreichte Reflexivitätsstufe mit $p = ,036$ einen signifikanten Unterschied zwischen beiden Faktorstufen, bei Messzeitpunkt 2 und 3 ist jedoch kein signifikanter Unterschied mehr feststellbar mit $p_2 = ,523$ und $p_3 = ,129$.

Ein Blick auf die Mittelwerte der Variablen zu Messzeitpunkt 2 zeigt, dass der Mittelwert des Gesamtscores im Reflektieren und des Reflexionsscores bei der Gruppe, die die Hütemethode als erste Teilintervention erhält, größer ist als bei der Gruppe, die Gibbs Reflexivitätszyklus als

erste Teilintervention bekommt. Beim Reflexivitätsscore sind die Mittelwerte der Gibbs Gruppe hingegen größer. Die am häufigsten erreichte Reflexivitätsstufe ist zu Messzeitpunkt 1 durchschnittlich ebenfalls bei der Gibbs Gruppe größer als bei der Hüte-Gruppe.

Nachdem festgestellt wurde, dass alle Variablen mit Ausnahme der Variable am häufigsten erreichte Reflexivitätsstufe vor den beiden Teilinterventionen keine signifikanten Unterschiede aufweisen und sich nach der ersten Teilintervention bei den Variablen Gesamtscore im Reflektieren, Reflexionsscore und Reflexivitätsscore signifikante Unterschiede zeigen, werden nun die Haupteffekte des Innersubjektfaktors Messzeitpunkt vorgestellt. Der Messzeitpunkt zeigt einen linearen, statistisch signifikanten Effekt auf folgende Variablen der Gibbs-Gruppe: Gesamtscore im Reflektieren ($F(21, 1) = 108,41, p < ,001, \text{partielles } \eta^2 = ,92$), Reflexionsscore ($F(21, 1) = 49,78, p < ,001, \text{partielles } \eta^2 = ,83$), Reflexivitätsscore ($F(21, 1) = 15,66, p = ,003, \text{partielles } \eta^2 = ,61$) und maximal erreichte Reflexivitätsstufe ($F(21, 1) = 11,27, p < ,007, \text{partielles } \eta^2 = ,53$). Auf die Variablen Reflexionsscore, Reflexivitätsscore und maximal erreichte Reflexivitätsstufe der Gibbs-Gruppe zeigen sich außerdem quadratisch statistisch signifikante Effekte von Zeit. Bei der Hüte-Gruppe sind auf folgende Variablen lineare, statistisch signifikante Effekte feststellbar: Gesamtscore im Reflektieren mit $p < ,001$ und partielles $\eta^2 = ,90$, Reflexivitätsscore mit $p = ,003$ und partielles $\eta^2 = ,57$, maximal erreichte Reflexivitätsstufe mit $p < ,001$ und partielles $\eta^2 = ,86$ und am häufigsten erreichte Reflexivitätsstufe mit $p = ,009$ und partielles $\eta^2 = ,48$. Quadratische, statistisch signifikante Effekte gibt es zusätzlich auf die Variable Gesamtscore im Reflektieren. Die Variable Reflexionsscore besitzt ausschließlich einen quadratischen, auffälligen Effekt hinsichtlich des Messzeitpunkts mit $F(1, 21) = 16,84, p = ,002$ und einem partiellen $\eta^2 = ,61$. Auf die Variable am häufigsten erreichte Reflexivitätsstufe der Gibbs-Gruppe lassen sich keine signifikanten Effekte hinsichtlich des Messzeitpunkts feststellen. Die Variablen maximal erreichte Reflexionsstufe und am häufigsten erreichte Reflexionsstufe, die keine Interaktionseffekte besitzen, sind lineare Haupteffekte hinsichtlich der Zeit feststellbar. Die maximal erreichte Reflexionsstufe hat dabei einen Wert von $F(1, 21) = 21,98$ und $p < ,001$, partielles $\eta^2 = ,51$ und die am häufigsten erreichte Reflexionsstufe von $F(1, 21) = 6,77$ und $p = ,017$, partielles $\eta^2 = ,24$.

Im nächsten Schritt wird geklärt, welche Messzeitpunkte genau innerhalb der einzelnen Faktorstufen signifikante Unterschiede bei den abhängigen Variablen haben. Für Hypothese 2, die prüft, ob sich die Gibbs-Gruppe hinsichtlich der Variablen mit Bezug auf die Reflexivitätsfähigkeit durch Gibbs Reflexivitätszyklus signifikant bessert, sind die Vergleiche der Variablen Reflexivitätsscore und maximal erreichte Reflexivitätsstufe der Messzeitpunkte 1 und 2 und 2 und 3 beider Experimentalgruppen von Bedeutung. Die Variable am häufigsten erreichte Reflexivitätsstufe wäre zwar hierfür relevant, allerdings ist hier kein Haupteffekt bezüglich der Messzeitpunkte feststellbar. Beide Variablen Reflexivitätsscore und maximal erreichte Reflexivitäts-

stufe unterscheiden sich signifikant zwischen Messzeitpunkt 1 und 2 mit $p = .004$ bei der Variablen Reflexivitätsscore (- 5,40, 95%-CI[- 8,86, - 1,94]) und mit $p = .008$ bei der Variablen maximal erreichte Reflexivitätsstufe (-,27, 95%-CI[- 1,02, - ,17]) der Experimentalgruppe 1. Die Gruppe, die als erste Teilintervention die Hütemethode erfährt, zeigt von Messzeitpunkt 1 auf 2 keine signifikanten Differenzen der entsprechenden Variablen, ebenso bei der Variable am häufigsten erreichte Reflexivitätsstufe. Von Messzeitpunkt 2 auf 3 sind bei Faktorstufe 1 keine signifikanten Unterschiede beider Variablen feststellbar ($p = .135$ beim Reflexivitätsscore und $p = .161$ bei der maximal erreichten Reflexivitätsstufe). Bei Experimentalgruppe 2, die Gibbs Reflexivitätszyklus als zweite Teilintervention erhält, verändern sich der Reflexivitätsscore von Messzeitpunkt 2 auf 3 mit $p < .001$ (- 3,33, 95%-CI[- 4,53, - 2,12]) und die Variable maximal erreichte Reflexivitätsstufe mit ebenfalls $p < .001$ (- ,63, 95%-CI[- ,93, - ,32]) deutlich signifikant.

Auf den Reflexivitätsscore hat die erste Teilintervention in Form von Gibbs Reflexivitätszyklus einen hohen Effekt von $d = 1,443$ ($M_1 = 18,88$; $M_2 = 22,62$) und auf die Variable maximal erreichte Reflexivitätsstufe einen hohen, etwas geringeren Effekt von $d = ,919$ ($M_1 = 2,54$; $M_2 = 3,08$). Gibbs als zweite Teilintervention hat ebenfalls einen hohen Effekt auf den Reflexivitäts-score mit $d = 1,400$ ($M_1 = 8,05$; $M_2 = 11,38$) und einen etwas geringeren, aber dennoch hohen Effekt auf die maximal erreichte Reflexivitätsstufe mit $d = ,988$ ($M_1 = 2,75$; $M_2 = 3,38$). Die Effektstärken beider Variablen durch Gibbs Reflexivitätszyklus als erste und zweite Teilintervention sind ähnlich groß und verdeutlichen, dass die Reflektierhilfe gewünschte Effekte bewirkt.

Für Hypothese 3, die die Wirkung der Hütemethode als Teilintervention auf die Variablen, die relevant für die Reflexionsfähigkeit sind, prüft, sind die Variablen Reflexionsscore, maximal erreichte Reflexionsstufe und am häufigsten erreichte Reflexionsstufe von Messzeitpunkt 1 und 2 und von Messzeitpunkt 2 auf 3 relevant. Der Reflexionsscore ist mit einer Differenz von - 5,47 (95%-CI[- 9,81, - 1,13]) von Messzeitpunkt 1 auf 2 signifikant mit $p = ,014$. Die Variable maximal erreichte Reflexionsstufe ist mit $p = .060$ nicht signifikant unterschiedlich, ebenso die Variable am häufigsten erreichte Reflexionsstufe mit einer Differenz von - ,542 (95%-CI[- 1,13, - ,046]) mit $p = ,074$. Die Gibbs Gruppe zeigt nach der ersten Teilintervention keine signifikanten Veränderungen aller drei Variablen. Von Messzeitpunkt 2 auf 3 ist bei Faktorstufe 2 (Hüte-Gruppe) ebenfalls eine signifikante Veränderung hinsichtlich der Variable Reflexionsscore feststellbar (2,89, 95%-CI[,35, 5,44]) mit $p = ,025$. Diese Veränderung geht jedoch in die andere Richtung und der Score nimmt von Messzeitpunkt 2 auf 3 ab. Die Variablen maximal erreichte Reflexionsstufe und am häufigsten erreichte Reflexionsstufe haben keinen signifikanten Unterschied mit jeweils $p = 1,000$. Auch bei Faktorstufe 1 (Gibbs-Gruppe) entwickelt sich der Reflexionsscore signifikant (- 6,41, 95%-CI[- 9,31, -3,51]) mit $p < ,001$, während die

Variablen maximal erreichte Reflexionsstufe und am häufigsten erreichte Reflexionsstufe unverändert bleiben. Die Veränderung dabei ist jedoch im Unterschied zu Experimentalgruppe 1 wieder positiv gerichtet.

Die Hütemethode als erste Teilintervention hat einen hohen Effekt von $d = 1,451$ ($M_1 = 12,18$; $M_2 = 17,64$) auf die Reflexionsscore (vgl. Tabelle 31). Die Effekte für die maximal erreichte Reflexionsstufe und die am häufigsten erreichte Reflexionsstufe werden nicht berechnet, da hier keine signifikanten Unterschiede feststellbar sind. Die Hütemethode als zweite Teilintervention hat einen Effekt von $d = 1,886$ ($M_1 = 10,63$; $M_2 = 17,05$) auf den Reflexionsscore, der höher ist als der Effekt der Hütemethode als erste Teilintervention. Auch hier wird die Effektstärke für die beiden anderen Variablen nicht berechnet, da keine Signifikanzen zwischen Messzeitpunkt 2 und 3 herrschen. Die Teilintervention Hütemethode hat sowohl in Form der ersten Teilintervention als auch in Form der zweiten Teilintervention einen positiven Effekt auf den Reflexionsscore.

Tabelle 31 Effektstärken durch Reflektiermethoden

	Cohens d			
	Gibbs Reflexivitätszyklus		Hütemethode	
	Erste Teilintervention	Zweite Teilintervention	Erste Teilintervention	Zweite Teilintervention
Reflexionsscore	/	-,705	1,451	1,886
Maximal erreichte Reflexionsstufe	/	/	/	/
Am häufigsten erreichte Reflexionsstufe	/	/	/	/
Reflexivitätsscore	1,443	1,400	/	/
Maximal erreichte Reflexivitätsstufe	,919	,988	/	/
Am häufigsten erreichte Reflexivitätsstufe	/	/	/	/

Die Hypothesen 2 und 3 werden folglich bestätigt und die Nullhypothesen abgelehnt. Zwar verbessern sich nicht alle Teilvariablen der Reflexions- und Reflexivitätsfähigkeit entsprechend, doch der Score zeigt stets eine Verbesserung nach der entsprechenden Reflektierhilfe. So bewirkt Gibbs Reflexivitätszyklus sowohl als erste als auch zweite Teilintervention eine Steigerung des Reflexivitätsscores und der maximal erreichten Reflexivitätsstufe und die Hütemethode eine Steigerung des Reflexionsscores. Bei Gibbs als zweite Teilintervention kommt es sogar zu einer Abnahme des Reflexionsscores, sodass dieser dadurch nicht gefördert wird. Die anderen Testpersonen, die die entsprechenden Reflektierhilfen nicht erhalten, weisen hinsichtlich der zu untersuchenden Variablen keine signifikanten Veränderungen auf.

3.5.3.3 Hypothese 4

Wissenschaftliche Hypothese 4:	Die Intervention führt zu einer Veränderung der Begriffsdefinition von Reflektieren.	
Statistisch:	H ₀ : Die Variablen fachliche Nähe und enthaltene Teilbereiche des Reflektierens unterscheiden sich nicht signifikant vor und nach der Intervention bei allen drei Faktorstufen.	H ₁ : Die Variablen fachlichen Nähe und enthaltenen Teilbereiche des Reflektierens unterscheiden sich signifikant vor und nach der Intervention bei beiden Experimentalgruppen, während sie sich bei der Kontrollgruppe nicht signifikant unterscheiden.

Es besteht eine lineare Interaktion zwischen dem Messzeitpunkt und den Gruppen bei der Variable fachliche Nähe mit $F(2, 36) = 13,25$, $p < ,001$ und partielles $\eta^2 = ,42$. Hinsichtlich der Variable Teilbereiche des Reflektierens ist keine Interaktion feststellbar mit $p = ,512$. Hier besteht jedoch ein Haupteffekt von Zeit auf die Variable mit $F(1, 36) = 2,002$, $p = ,010$ und partielles $\eta^2 = ,17$. Ein Haupteffekt von Gruppe auf die entsprechende Variable ist nicht feststellbar mit $F(2, 36) = 2,61$, $p = ,088$ und partielles $\eta^2 = ,13$.

Nachfolgend werden nun die einfachen Haupteffekte des Zwischensubjektfaktors und des Innersubjektfaktors für die Variable fachliche Nähe untersucht. Zwischen den Gruppen unterscheidet sich die Variable fachliche Nähe weder vor der Intervention noch nach der Intervention signifikant mit ($p_1 = ,520$ und $p_2 = ,077$). Aufgrund des Nichtvorhandenseins eines einfachen Haupteffekts für die Variable fachliche Nähe und eines Haupteffekts für die Variable Teilbereiche des Reflektierens des Zwischensubjektfaktors sind post-hoc Test überflüssig.

Als nächster Schritt wird untersucht, ob einfache Haupteffekte des Innersubjektfaktors bezüglich der Variable fachliche Nähe der Definition bestehen. Es gibt statistisch signifikante Effekte von Zeit auf die Ausprägung der fachlichen Nähe der beiden Experimentalgruppen mit $F(1, 11) = 17,47$, $p = ,002$ und partielles $\eta^2 = ,61$ bei der Gibbs-Gruppe und $F(1, 11) = 55,00$, $p < ,001$ und partielles $\eta^2 = ,83$ bei der Hüte-Gruppe. Bei der Kontrollgruppe ist kein einfacher Haupteffekt hinsichtlich des Innersubjektfaktors vorhanden mit $p = 1,000$.

Anhand der paarweisen Vergleiche lässt sich erkennen, dass die Entwicklungen der Variable fachliche Nähe positiv gerichtet sind und die fachliche Nähe bei beiden Experimentalgruppen nach der Intervention signifikant größer ist als vor der Intervention. Für die Variable Teilbereiche des Reflektierens, für die Haupteffekte bezüglich der Zeit festgestellt wurden, gibt es keine signifikanten Unterschiede von Messzeitpunkt 1 auf 2 bei allen drei Faktorstufen mit $p = ,096$ der Experimentalgruppe 1, $p = ,137$ der Experimentalgruppe 2 und $p = ,334$ der Kontrollgruppe. Die Intervention hat auf die fachliche Nähe in den Begriffsdefinitionen einen Effekt von $d = 1,643$ bei Experimentalgruppe 1 und einen Effekt von $d = 1,454$ bei Experimentalgruppe

2. Beide Gruppen zeigen folglich hohe, ähnliche Effekte, bei der Kontrollgruppe sind keine Effekte berechenbar, da es keine signifikanten Unterschiede gibt.

Die Hypothese kann folglich nur zum Teil bejaht werden. Während sich die Anzahl der enthaltenen Teilbereiche des Reflektierens innerhalb der Definitionen vor und nach der Intervention aller Gruppen nicht signifikant unterscheiden, gibt es signifikante Unterschiede hinsichtlich der fachlichen Nähe. Diese verbessert sich signifikant bei beiden Experimentalgruppen und bleibt bei der Kontrollgruppe konstant. Die Hypothese kann bezüglich der Variable fachliche Nähe als bestätigt betrachtet werden und für diesen Bereich die Nullhypothese abgelehnt werden.

3.5.3.4 Hypothesengenerierender Teil

Die Studie liefert Daten, die zur Generierung neuer Hypothesen dienen. Dabei soll der Einfluss der Reihenfolge des Erlernens der Reflektiermethoden auf die Reflektierfähigkeiten näher untersucht werden, außerdem soll der Frage nachgegangen werden, inwieweit das Geschlecht und die persönliche Bewertung der Reflektierhilfen Einfluss auf die Entwicklung der Reflektierfähigkeiten nimmt. Weiterhin wird der Frage nachgegangen, ob und falls ja, inwieweit sich die persönlichen Einstellungen zum Reflektieren durch die Intervention verändern.

(1) Reihenfolge des Erlernens der Reflektiermethoden

Zunächst soll geklärt werden, ob eine bestimmte Reihenfolge des Erlernens der Reflektiermethoden einen signifikant größeren Effekt auf die Reflektierfähigkeit bestehend aus den einzelnen Variablen hat. Für die Generierung einer Hypothese werden zunächst die Reflektierfähigkeiten beider Gruppen des ersten und letzten Beitrags miteinander verglichen, anschließend die des ersten und letzten Textblocks.

Der Vergleich beider Gruppen des ersten und letzten Beitrags wurde bereits in Kapitel 3.5.3.1 behandelt, weshalb sich hier direkt auf die Effektstärken fokussiert wird (vgl. Tabelle 32). Der Effekt der Gesamtintervention auf den Gesamtscore im Reflektieren ist bei beiden Faktorstufen stark mit $d = 4,061$ bei Experimentalgruppe 1 ($M_1 = 19,02$ und $M_2 = 28,13$) und $d = 4,772$ bei Experimentalgruppe 2 ($M_1 = 19,18$ und $M_2 = 30,31$). Bei Experimentalgruppe 2, die die Reihenfolge Hüte – Gibbs haben, ist der Effekt auf den Reflexionsscore, der maximal erreichten Reflexionsstufe, der maximal erreichten Reflexivitätsstufe, der am häufigsten erreichten Reflexionsstufe und der am häufigsten erreichten Reflexivitätsstufe höher als bei Experimentalgruppe 1 (für die maximal und am häufigsten erreichte Reflexivitätsstufe gibt es keine Effektstärke, da keine signifikanten Unterschiede der Werte existieren). Auf die Variable Reflexivitätsscore gibt es einen Effekt bei Experimentalgruppe 1, bei Experimentalgruppe 2 ist keiner vorhanden. Die Reihenfolge Hüte - Gibbs weist zusammengefasst höhere Effekte des Reflektierens auf als die Reihenfolge Gibbs - Hüte, nur beim Reflexivitätsscore ist dies nicht der Fall.

Tabelle 32 Effektstärken nach d Cohen vom ersten zum letzten Beitrag beider Experimentalgruppen

Variable	d Cohen Experimental- gruppe 1 (Reihenfolge Gibbs – Hüte)	d Cohen Experimental- gruppe 2 (Reihenfolge Hüte – Gibbs)
Gesamtscore im Reflektieren	4,061	4,772
Reflexionsscore	1,442	1,792
Reflexivitätsscore	,617	/
Maximal erreichte Reflexionsstufe	1,459	1,745
Maximal erreichte Reflexivitätsstufe	/	,733
Am häufigsten erreichte Reflexions- stufe	1,051	1,333
Am häufigsten erreichte Reflexivitäts- stufe	/	1,405

Als zweiter Schritt werden nun die Reflektierentwicklungen beider Gruppen vom ersten zum letzten Textblock herangezogen. Dafür werden alle Variablen, die (einfache) Haupteffekte des Innersubjektfaktors aufweisen, näher betrachtet. Die Variable am häufigsten erreichte Reflexivitätsstufe ist deshalb nur für Experimentalgruppe 2 relevant. Bei Experimentalgruppe 1, die Gibbs Reflexivitätszyklus zuerst erhält, weisen alle Variablen mit Ausnahme der Variable am häufigsten erreichte Reflexionsstufe ($p = ,159$) signifikante Differenzen zwischen Messzeitpunkt 1 und 3 auf: Gesamtscore im Reflektieren mit $p < ,001$ ($- 7,56$, 95%-CI[$- 9,65$, $- 5,48$]), Reflexionsscore mit $p < ,001$ ($- 5,00$, 95%-CI[$- 7,03$, $- 2,97$]), Reflexivitätsscore mit $p = ,004$ ($- 2,59$, 95%-CI[$- 4,47$, $- ,71$]), maximal erreichte Reflexionsstufe mit $p = ,016$ ($- ,64$, 95%-CI[$- 1,15$, $- ,12$]) und maximal erreichte Reflexivitätsstufe mit $p = ,022$ ($- ,59$, 95%-CI[$- 1,10$, $- ,09$]). Auch bei Experimentalgruppe 2, die die Hütemethode zuerst erhält, sind bei den meisten Variablen signifikante Unterschiede zwischen Messzeitpunkt 1 und 3 feststellbar: Gesamtscore im Reflektieren mit $p < ,001$ ($- 6,81$, 95%-CI[$- 8,79$, $- 4,83$]), Reflexivitätsscore mit $p = ,008$ ($- 4,23$, 95%-CI[$- 7,34$, $- 1,17$]), maximal erreichte Reflexionsstufe mit $p = ,036$ ($- ,38$, 95%-CI[$- ,73$, $- ,02$]), maximal erreichte Reflexivitätsstufe mit $p < ,001$ ($- 1,00$, 95%-CI[$- 1,35$, $- ,65$]) und am häufigsten erreichte Reflexivitätsstufe mit $p = ,027$ ($- ,79$, 95%-CI[$- 1,50$, $- ,09$]). Die Variable Reflexionsscore bleibt hinsichtlich signifikanter Unterschiede konstant mit $p = ,335$.

Bei den Testpersonen, die erst Gibbs Reflexivitätszyklus und anschließend die Hütemethode erhalten, zeigt die Intervention hinsichtlich des Gesamtscores im Reflektieren eine sehr hohe Effektstärke von $d = 3,217$ ($M_1 = 18,98$ und $M_2 = 26,44$; vgl. Tabelle 33). Auf den Reflexionsscore hat die Intervention einen Effekt von $d = 1,516$ ($M_1 = 11,94$ und $M_2 = 17,05$), auf den Reflexivitätsscore von $d = ,923$ ($M_1 = 7,02$ und $M_2 = 9,39$), auf die maximal erreichte Reflexionsstufe von $d = 1,152$ ($M_1 = 3,13$ und $M_2 = 3,71$) und auf die maximal erreichte Reflexivitätsstufe von $d = ,948$ ($M_1 = 2,54$ und $M_2 = 3,04$). Bei den Testpersonen, die die Reflektierhilfen in der Reihenfolge Hütemethode - Gibbs erhalten, sind folgende Effekte der Intervention auf die

Variablen feststellbar: $d = 3,362$ beim Gesamtscore im Reflektieren ($M_1 = 19,28$ und $M_2 = 25,67$), $d = 1,449$ beim Reflexivitätsscore ($M_1 = 7,12$ und $M_2 = 11,38$), $d = ,687$ bei der maximal erreichten Reflexionsstufe ($M_1 = 3,13$ und $M_2 = 3,5$), $d = 2,181$ bei der maximal erreichten Reflexivitätsstufe ($M_1 = 2,38$ und $M_2 = 3,38$) und $d = 1,121$ bei der am häufigsten erreichten Reflexivitätsstufe ($M_1 = 1,63$ und $M_2 = 2,42$).

Tabelle 33 Effektstärken nach d Cohen vom ersten zum letzten Textblock beider Experimentalgruppen

Variable	d Cohen Experimental- gruppe 1 (Reihenfolge Gibbs - Hüte)	d Cohen Experimental- gruppe 2 (Reihenfolge Hüte - Gibbs)
Gesamtscore im Reflektieren	3,217	3,362
Reflexionsscore	1,516	/
Reflexivitätsscore	,923	1,449
Maximal erreichte Reflexionsstufe	1,152	,687
Maximal erreichte Reflexivitätsstufe	,948	2,181
Am häufigsten erreichte Reflexions- stufe	/	/
Am häufigsten erreichte Reflexivitäts- stufe	/	1,121

Zusammengefasst kann festgehalten werden, dass sich der Gesamtscore im Reflektieren vom ersten Textblock zum letzten Textblock bei der Gruppe, die zunächst die Hütemethode als Teilintervention erhält, leicht stärker verbessert als bei der Gruppe, die zunächst Gibbs Reflexivitätszyklus als Teilintervention erhält. Auch der Reflexivitätsscore, die maximal erreichte Reflexivitätsstufe und die am häufigsten erreichte Reflexivitätsstufe nehmen bei dieser Gruppe deutlich mehr zu. Der Reflexionsscore und die maximal erreichte Reflexionsstufe erhöhen sich jedoch bei der Gruppe, die Gibbs zuerst bekommt, stärker als bei der Hüte-Gruppe.

Aus beiden Ergebnissen kann die Hypothese abgeleitet werden, dass die Reihenfolge der Reflektierhilfen Einfluss auf die Entwicklung der Reflektierfähigkeit nimmt und die Reihenfolge Hütemethode – Gibbs Reflexivitätszyklus als sinnvoller erscheint. Weiterhin kann die Hypothese generiert werden, dass die Reflektiermethode, die zuletzt erhalten wird, stärkeren Einfluss auf den zu fördernden Teilbereich des Reflektierens nimmt und dieser folglich höher ausgeprägt ist als der andere Teilbereich, der zuvor mehr gefördert wird. Dies wird daraus abgeleitet, dass die Entwicklung von Textblock 1 auf 3 anders ausfallen als die Entwicklung vom ersten zum letzten Beitrag.

(2) Geschlecht

Zwischen dem Zwischensubjektfaktor Geschlecht und dem Innersubjektfaktor Zeit gibt es bei den Variablen Reflexionsscore und Reflexivitätsscore signifikante Interaktionseffekte mit $F(1, 22) = 5,26$ und $p = ,032$ der ersten Variable und $F(1, 22) = 4,84$ und $p = ,039$ der zweiten Variable. Die übrigen Variablen, die die Reflektierfähigkeit näher definieren, zeigen zwischen den beiden Faktoren keine Interaktionseffekte.

Für den Reflexionsscore und den Reflexivitätsscore werden die einfachen Innersubjekt- und Zwischensubjekteffekte näher betrachtet. Bei beiden ist kein signifikanter Unterschied hinsichtlich des Zwischensubjektfaktors Gruppe feststellbar. Die Variable Reflexionsscore zum ersten Messzeitpunkt hat einen Signifikanzwert von $p = ,497$ und beim zweiten Messzeitpunkt von $p = ,132$. Die Variable Reflexivitätsscore hat zum ersten Messzeitpunkt einen Signifikanzwert von $p = ,481$ und von $p = ,175$ beim zweiten Messzeitpunkt. Für die restlichen Variablen, die keinen Interaktionseffekt Zeit x Gruppe aufweisen, werden die Haupteffekte bestimmt. Hinsichtlich der Gruppe liegt ein signifikanter Haupteffekt lediglich bei der Variablen am häufigsten erreichte Reflexionsstufe vor mit $F(1, 22) = 8,88$ und $p = ,007$.

Die Variablen maximal erreichte Reflexionsstufe und am häufigsten erreichte Reflexionsstufe des letzten Messzeitpunktes zeigen signifikante Unterschiede zwischen den beiden Faktorstufen Weiblich und Männlich mit $p = ,031$ und $p = ,017$. Dabei sind die durchschnittlichen Werte der Faktorstufe Männlich bei beiden Variablen höher. Alle anderen Variablen des ersten und letzten Messzeitpunkts weisen keine signifikanten Unterschiede auf. Beide Faktorstufen haben folglich gleiche Ausgangsvoraussetzungen, sodass ihre Entwicklungen miteinander vergleichbar sind.

Die Zeit hat einen signifikanten Einfluss auf beide Variablen Reflexionsscore und Reflexivitätsscore mit $F(1, 9) = 13,95$ und $p = ,005$ und $F(1, 9) = 11,71$ der Faktorstufe weiblich. Bei der Faktorstufe männlich ist nur ein signifikanter Effekt von Zeit auf den Reflexionsscore mit $F(1, 13) = 49,63$ und $p < ,001$ vorhanden. Bei allen drei Variablenentwicklungen mit signifikanten Unterschieden über den Zeitraum hinweg, verändern sich die Werte positiv, d.h. durch die Intervention steigen diese an. Bei allen Variablen tritt ein signifikanter Haupteffekt Zeit auf mit $p < ,001$ mit Ausnahme des Reflexivitätsscores ($p = ,001$).

Auf allen Variablen mit Ausnahme der Variable am häufigsten erreichte Reflexionsstufe gibt es bei der Faktorstufe Weiblich signifikante Effekte von Zeit. Auch bei der Faktorstufe Männlich hat die Zeit signifikante Bedeutung für die meisten Variablen mit Ausnahme der Variablen Reflexivitätsscore und am häufigsten erreichte Reflexivitätsstufe mit $p = ,225$ und $p = ,144$. Die paarweisen Vergleiche zeigen, dass sich alle Variablen mit signifikanten Veränderungen durch die Intervention positiv verändern.

Die Effektstärken werden für diejenigen Variablen bestimmt, die signifikante Unterschiede durch den Faktor Zeit innerhalb der beiden Faktoren Geschlecht haben (vgl. Tabelle 34). Die

Intervention hat auf den Gesamtscore im Reflektieren bei männlichen Testpersonen einen leicht höheren Effekt als bei den weiblichen mit $d = 4,353$ ($M_1 = 19,09$ und $M_2 = 29,43$) und $d = 3,952$ ($M_1 = 19,11$ und $M_2 = 28,92$). Dies ist ebenfalls bei den Variablen Reflexionsscore und maximal erreichte Reflexionsstufe der Fall. Bei den Variablen Reflexivitätsscore und am häufigsten erreichte Reflexivitätsstufe sind nur Effekte der Intervention auf die weiblichen Testpersonen feststellbar mit $d = 1,119$ ($M_1 = 6,82$ und $M_2 = 11,86$) und $d = 1,692$ ($M_1 = 1,70$ und $M_2 = 2,20$). Auch bei der maximal erreichten Reflexivitätsstufe sind die Effekte bei der Faktorstufe weiblich höher als bei der männlichen. Die am häufigsten erreichte Reflexionsstufe verändert sich nur bei den männlichen Testpersonen signifikant mit einer Effektstärke von $d = 1,542$ ($M_1 = 2,00$ und $M_2 = 3,04$).

Tabelle 34 Effektstärken hinsichtlich des Geschlechts

Variable	d Cohen Faktorstufe weiblich	d Cohen Faktorstufe männlich
Gesamtscore im Reflektieren	3,952	4,353
Reflexionsscore	1,108	2,071
Reflexivitätsscore	1,119	/
Maximal erreichte Reflexionsstufe	1,226	1,993
Maximal erreichte Reflexivitätsstufe	1,606	,581
Am häufigsten erreichte Reflexionsstufe	/	1,542
Am häufigsten erreichte Reflexivitätsstufe	1,692	/

Aus den Ergebnissen können folgende Hypothesen entwickelt werden:

- Die Intervention hat auf weibliche Teilnehmende einen höheren, positiven Einfluss auf die Reflexivitätsfähigkeit als auf Männliche.
- Die Intervention hat auf männliche Teilnehmende einen höheren, positiven Einfluss auf die Reflexionsfähigkeiten als auf Weibliche.
- Die Intervention hat auf beide Geschlechter ein hohes Potenzial für die Verbesserung der Reflektierfähigkeiten insgesamt.

(3) Persönliche Bewertung der Reflektiermethoden

Die Verteilung der persönlichen Einschätzungen zu beiden Reflektiermethoden variiert sehr stark, sodass die beiden Stichproben für die Untersuchung des Einflusses der persönlichen Bewertung der Reflektierhilfen zu unterschiedlich sind, um verlässliche Ergebnisse zu liefern.³⁴ Deshalb kann dieser Bereich nicht statistisch untersucht werden.

³⁴ Lediglich drei Testpersonen bewerten die Hütemethode als nicht hilfreich und fünf Testpersonen geben an, dass Gibbs Reflexivitätszyklus nicht hilfreich sei. Die übrigen 21 und 19 Testpersonen bewerten sie jeweils als hilfreich.

(4) Persönliche Einstellung zum Reflektieren

Es gibt keine signifikante Interaktion zwischen der Zeit und den drei Untersuchungsgruppen (Experimentalgruppe 1 und 2 und Kontrollgruppe) hinsichtlich der Variablen der persönlichen Einstellungen zum Reflektieren. Die Signifikanzen der Variablen Entscheidungen, Erklärung, Gründe, Akzeptanz, Andere und Nachdenken vom ersten zum letzten Messzeitpunkt liegen alle im Bereich $,150 < p > ,989$. Auch lässt sich kein Haupteffekt von Zeit auf die Angaben der persönlichen Einstellungen zum Reflektieren hinsichtlich der sechs Variablen feststellen.

Es existiert auch kein Haupteffekt vom Zwischensubjektfaktor Gruppe, die Signifikanzen variieren hierbei bei den sechs Variablen Entscheidung, Erklärung, Gründe, Akzeptanz, Andere und Nachdenken zwischen $p = ,173$ und $p = ,969$.

Die Angaben bleiben folglich durch die Intervention unverändert, genauso wie bei der Kontrollgruppe. Somit kann die Hypothese im Rahmen dieser Untersuchung generiert werden, dass die Intervention keinen Einfluss auf die abgegebenen persönlichen Einstellungen zum Reflektieren nimmt.

4 Beantwortung der Forschungsfragen

Für die Beantwortung der Forschungsfrage werden nachfolgend zunächst die Teilforschungsfragen bearbeitet und sich anschließend der Hauptforschungsfrage gewidmet. Die Antworten setzen sich aus den Ergebnissen der qualitativen Auswertung und der nachgelagerten quantitativen Auswertung zusammen. Die ersten vier Teilforschungsfragen enthalten Ergebnisse beider Auswertungsschritte. Ab der fünften Frage wird sich für die Beantwortung lediglich an den quantitativen Ergebnissen orientiert. Die Kombination beider Auswertungstechniken gewährleistet eine ganzheitliche, multiperspektivische Untersuchung der Forschungsfragen.

(1) Welches Potenzial bieten die beiden Reflektiermethoden insgesamt für die Entwicklung der Reflektierfähigkeiten beim Bloggen?

Im Rahmen der qualitativen Auswertung mittels der qualitativen Inhaltsanalyse nach Mayring (2015) zeigt sich, dass alle Testpersonen der Experimentalgruppe, die beide Reflektiermethoden erlernen, insgesamt positive Entwicklungen der Reflektierfähigkeit vom ersten zum letzten Beitrag aufweisen. Sieben Schülerinnen und Schüler verbessern sich hinsichtlich aller Kriterien, die für die Bestimmung der Reflektierfähigkeit definiert wurden.³⁵ Elf Probandinnen und Probanden zeigen neben Verbesserungen der Kriterien auch sich nicht verändernde Werte und bei vier Testpersonen gibt es bei einem Kriterium eine negative Entwicklung. Nur eine Testperson zeigt mehr als eine negative Veränderung hinsichtlich der sieben Kriterien. Dies ist damit zu begründen, dass der letzte Beitrag dieser Person (P4) keine Reflexivitätsanteile enthält und folglich alle Kriterien, die die Reflexivität betreffen, null ergeben. Insgesamt betrachtet ist also eine überwiegend verbessernde Tendenz feststellbar. Der Gesamtscore im Reflektieren steigt bei allen Testpersonen mit Zunahmen zwischen den Werten 6,0 und 15,5 an. Alle Gesamtscores liegen nach der Intervention höher als der maximal gemessene Gesamtscore vor der Intervention, mit Werten über 23,8. Auch die Reflexions- und Reflexivitätsscores erhöhen sich bei allen mit Ausnahme jeweils einer Person mit einer Verschlechterung und zwei Personen mit konstant bleibenden Werten beim Reflexivitätsscore. Die Zunahmen liegen beim Reflexionsscore zwischen den Werten 1,2 und 16,5 und beim Reflexivitätsscore zwischen ,5 und 17,5. Bei der maximal und am häufigsten erreichten Reflexionsstufe gibt es ebenfalls nur positive Veränderungen bzw. gleichbleibende Werte (zwischen 0,0 und 3,0 und 0,0 und 2,0), bei der maximal erreichten Reflexivitätsstufe zeigt sich bei nur einer Testperson eine negative Tendenz (Zunahmen zwischen den Werten 0,0 und 2,0). Bei der am häufigsten erreichten Reflexivitätsstufe haben zwei Teilnehmende negative Veränderungen (Zunahmen zwischen den Werten 0,0 und 2,5).

³⁵ Zu den Kriterien zählen der Gesamtscore im Reflektieren, der sich aus einem Reflexionsscore und einem Reflexivitätsscore zusammensetzt, die maximal erreichten Reflexions- und Reflexivitätsstufen und die am häufigsten erreichten Reflexions- und Reflexivitätsstufen.

Die quantitative Auswertung verifiziert die aufgestellte Hypothese, dass die Intervention zu einer gesteigerten Reflektierfähigkeit führe. Während sich die einzelnen Variablen der Reflektierfähigkeit³⁶ bei der Kontrollgruppe nicht signifikant verändern, sind signifikante Veränderungen der meisten Variablen der beiden Experimentalgruppen, die sich hinsichtlich der Reihenfolge der erhaltenen Reflektierhilfen unterscheiden, vorhanden. Alle Effekte fallen dabei nach der Interpretation nach Cohen (1988) groß und positiv aus. Lediglich bei drei Variablen sind nicht in beiden Experimentalgruppen signifikante Veränderungen über die Intervention hinweg feststellbar (maximal erreichte Reflexivitätsstufe, am häufigsten erreichte Reflexivitätsstufe, Reflexivitätsscore). Die Variablen zeigen vor der Intervention bei allen drei Gruppen keine signifikanten Unterschiede, sodass von gleichen Voraussetzungen ausgegangen werden kann und die Entwicklungen miteinander verglichen werden können. Nach der Intervention unterscheiden sich die Variablen zwischen der Kontrollgruppe und den Experimentalgruppen signifikant. Auf beide Experimentalgruppen hat die Intervention positive Wirkungen. Keine der beiden Gruppen zeigt bei einer Variable negative Durchschnittsdifferenzen zwischen dem letzten und den ersten Beitrag.

Fasst man die Ergebnisse der qualitativen und quantitativen Auswertung für die Beantwortung der ersten Frage zusammen, kommt man zu dem Ergebnis, dass die Intervention insgesamt positive Effekte auf die Reflektierfähigkeit erzielt. Die Testpersonen reflektieren vor der Intervention über Themen überwiegend auf dem Niveau Bezug zu nehmen, d.h. sie äußern ihre Meinungen zu den Themen, ohne diese zu begründen oder näher zu untersuchen. Nach der Intervention ist das Niveau Herstellung von Zusammenhängen am häufigsten vorhanden. Die Probandinnen und Probanden sind am Ende des Projekts vermehrt in der Lage, auch Gründe für Sacherhalte und Ansichten zu nennen. Hinsichtlich der Reflexivität werden bereits vor der Intervention hauptsächlich Gründe für Verhaltensweisen genannt. Das am häufigsten erreichte Reflexivitätsniveau ist nach der Intervention die analytische Abstraktion, bei der die Testpersonen unterschiedliche Perspektiven berücksichtigen und Gründe zueinander in Beziehung setzen. Die Antwort auf die erste Teilforschungsfrage lautet folglich: Die Wirkung der Intervention ist sehr groß auf die Reflektierfähigkeit von Schülerinnen und Schülern beim Bloggen und zeigt sich bei beiden Teilbereichen des Reflektierens, der Reflexion und der Reflexivität.

³⁶ Für die quantitative Auswertung wird der Begriff *Variablen* für die Kriterien verwendet, die die Reflektierfähigkeit bestimmen.

(2) Welches Potenzial bietet die Reflektiermethoden in Form von Gibbs Reflexivitätszyklus für die Entwicklung der Reflektierfähigkeiten?

Die qualitative Auswertung kommt zu dem Ergebnis, dass Gibbs Reflexivitätszyklus als Teilintervention unterschiedliche Potenziale für die Probandinnen und Probanden bietet. Tendenziell kann eine Verbesserung der Reflexivitätsfähigkeit festgestellt werden, während die Reflexionsfähigkeit überwiegend auf einem ähnlichen Niveau verbleibt. Im Detail zeigen die Ergebnisse, dass sich nach Gibbs Reflexivitätszyklus als erste Teilintervention eine Testperson hinsichtlich aller drei Reflexivitätskriterien – dem Reflexivitätsscore, der maximal erreichten Reflexivitätsstufe und der am häufigsten erreichten Reflexivitätsstufe – positiv entwickelt. Sieben Teilnehmende zeigen sowohl positive als auch keine Entwicklungen bei den drei Reflexivitätskriterien. Drei der insgesamt elf Testpersonen der Experimentalgruppe 1 weisen auch mindestens eine negative Entwicklung auf. Alle drei Kriterien, die die Reflexionsfähigkeit bestimmen, verbessern sich nur bei einer Person nach Gibbs Reflexivitätszyklus als erste Teilintervention. Sieben Schülerinnen und Schüler zeigen überwiegend keine Veränderungen hinsichtlich der drei Kriterien und drei Testpersonen weisen auch negative Entwicklungen bezüglich der Reflexionsfähigkeit auf. Gibbs Reflexivitätszyklus als zweite Teilintervention bewirkt ähnliche, leicht bessere Veränderungen der Reflexivitätskriterien gegenüber der Reflektierhilfe als erste Teilintervention: Fünf Teilnehmende haben nur positive Entwicklungen hinsichtlich der drei Kriterien und sieben Jugendliche positive und keine Entwicklungen. Keiner der Testpersonen weist bei einem Reflexivitätskriterium eine negative Entwicklung auf. Die Reflexionsfähigkeit dieser entwickelt sich dabei hinsichtlich mindestens eines Kriteriums negativ bei allen Schülerinnen und Schülern mit Ausnahme einer Testperson, die bei allen drei Kriterien Verbesserungen zeigt.

Die Ergebnisse der statistischen Auswertung bestätigen die Hypothese, dass Gibbs Reflexivitätszyklus als erste und zweite Teilintervention Veränderungen der Reflexivitätsfähigkeit bewirkt, während sich die Reflexionsfähigkeit nicht signifikant verändert. Dieser positiv gerichtete Effekt lässt sich dabei jedoch nur beim Reflexivitätsscore und der maximal erreichten Reflexivitätsstufe feststellen, da sich die am häufigsten erreichte Reflexivitätsstufe nicht signifikant zwischen den Messzeitpunkten unterscheidet. Gibbs Reflexivitätszyklus als erste Teilintervention hat einen Effekt mit $d = 1,443$ auf den Reflexivitätsscore und mit $d = ,919$ auf die maximale erreichte Reflexivitätsstufe. Als zweite Teilintervention betragen die Effekte $d = 1,400$ und $d = ,988$ und sind somit fast gleich wie zuvor.

Die Forschungsfrage kann mithilfe der Erkenntnisse aus beiden Auswertungsverfahren wie folgt beantwortet werden: Gibbs Reflexivitätszyklus ist eine Methode, die vor allem zur Förderung der Reflexivitätsfähigkeit angewendet werden kann. Die meisten Testpersonen steigern ihre Reflexivitätsfähigkeit nachdem sie mit Gibbs Reflexivitätszyklus vertraut gemacht wurden. Auch aus statistischer Sicht sind positive Effekte auf den Reflexivitätsscore und der maximal

erreichten Reflexionsstufe feststellbar. Während die Probandinnen und Probanden vor Gibbs Reflexivitätszyklus überwiegend über Handlungen reflektieren, indem sie eigenes Verhalten begründen, aber nicht genauer analysieren, finden nach Kenntnis der Methode häufiger Analysen von Verhalten durch die Beachtung unterschiedlicher Perspektiven statt. Für die Förderung der Reflexionsfähigkeit bietet Gibbs kein erkennbares Potenzial, da sie sich nach der Teilintervention nicht signifikant verändert und die Werte der einzelnen Testpersonen überwiegend gleichbleiben und nur vereinzelt positive Entwicklungen vorhanden sind. Das Potenzial von Gibbs Reflexivitätszyklus hinsichtlich der Förderung der Reflexivität als Teilbereich des Reflektierens, das bereits anhand der Literatur vermutet wurde, wird folglich durch diese Studie bestätigt.

(3) Welches Potenzial bietet die Reflektiermethode in Form der Hütemethode für die Entwicklung der Reflektierfähigkeiten?

Die qualitative Auswertung der Entwicklungen der Reflektierfähigkeit durch die Hütemethode nach de Bono zeigt, dass alle Proben mit Ausnahme von drei Testpersonen nur positive und keine Entwicklungen des Reflexionsscores, der maximal erreichten Reflexionsstufe und der am häufigsten erreichten Reflexionsstufe nach Kenntnis der Hütemethode aufweisen. In Form der ersten Teilintervention ändern sich die Reflexionskriterien bei den meisten positiv mit mindestens einem konstant bleibenden Wert. Drei Schülerinnen und Schüler besitzen nur positive Veränderungen. Eine Testperson weist zusätzlich eine negative Veränderung bei einer der drei Kriterien auf. Auch als zweite Teilintervention sind die Entwicklungen ähnlich: Die meisten Probandinnen und Probanden haben positive und keine Veränderungen hinsichtlich der Reflexionskriterien und jeweils zwei Testpersonen zeigen nur positive oder zusätzliche negative Entwicklungen. Während sich die Reflexionsfähigkeit überwiegend positiv durch die Hütemethode verändert, sind bei der Reflexivitätsfähigkeit andere Entwicklungen feststellbar: Bei sieben Teilnehmenden gibt es mindestens eine negative Veränderung der Reflexivitätskriterien (Score, maximal erreichte und am häufigsten erreichte Stufe) nach der Hütemethode als erste Teilintervention. Zwei Testpersonen haben nur positive Veränderungen und drei haben neben positiven Entwicklungen mindestens einmal keine Veränderung. Auch als zweite Teilintervention hat die Hütemethode keine erkennbare Optimierung der Reflexivitätsfähigkeiten. Nur eine Testperson weist bei allen drei Reflexivitätskriterien positive Entwicklungen auf und eine weitere Person neben positiven auch keine Entwicklungen. Alle übrigen Probandinnen und Probanden besitzen mindestens eine negative Entwicklung bei den drei Kriterien und die Reflexivitätsfähigkeit insgesamt nimmt eher ab als zu.

Die quantitative Auswertung bestätigt die Beobachtungen der qualitativen Auswertung für die Entwicklung der Reflexions- und Reflexivitätsfähigkeit durch die Hütemethode. Während sich der Reflexionsscore mit Effektstärken von $d = 1,451$ bei der Hütemethode als erste Teilintervention und $d = 1,886$ als zweite Teilintervention statistisch gesehen signifikant verbessert,

bleiben die drei Kriterien für die Reflexivitätsfähigkeit signifikant unverändert bzw. der Reflexivitätsscore verschlechtert sich sogar bei letzterem. Die maximal erreichte Reflexionsstufe und die maximal erreichte Reflexivitätsstufe verändern sich auch bei Erhalt der Hütemethode nicht signifikant.

Insgesamt kann die Teilforschungsfrage folgendermaßen beantwortet werden: Das Potenzial der Hütemethode für die Entwicklung der Reflektierfähigkeit entfaltet sich beim Teilbereich Reflexion. Die meisten Testpersonen zeigen hinsichtlich des Scores, der maximal erreichten und am häufigsten erreichten Reflexionsstufe Verbesserungen nach Erhalt der beschriebenen Reflektierhilfe. Auch statistisch gesehen nimmt der Reflexionsscore dabei signifikant zu. Die anderen Teilnehmenden verbessern ihre Reflexivitätsfähigkeit überwiegend nicht durch die Hütemethode, sondern zeigen gleichbleibende oder sogar negative Entwicklungen hinsichtlich des Reflexivitätsscores, der maximal erreichten und der am häufigsten erreichten Reflexivitätsstufe. Die Annahme, die Hütemethode führe zu einer Verbesserung des Reflexionsscores, kann folglich bestätigt werden und stützt somit in der Literatur beschriebene Potenziale der Hütemethode von de Bono.

(4) Inwieweit ändert sich das Begriffsverständnis von Reflektieren durch das Blogprojekt?

Die durch die Testpersonen abgegebenen Definitionen zum Begriff *Reflektieren* wurden anhand der fachlichen Nähe und der Anzahl der enthaltenen Teilbereiche des Reflektierens genauer untersucht. Die qualitative Auswertung der Begriffsdefinitionen hinsichtlich der fachlichen Nähe und der enthaltenen Teilbereiche des Reflektierens in der Definition liefern das Ergebnis, dass die meisten Schülerinnen und Schüler ein besseres Verständnis von Reflektieren nach der Intervention aufweisen als zuvor. 18 von 24 Teilnehmenden steigern die fachliche Nähe ihrer Definitionen und die meisten Definitionen weisen nach der Intervention eine hohe fachliche Nähe auf. Die anderen sechs Personen zeigen keine Veränderung hinsichtlich der fachlichen Nähe, sodass sich folglich keine Testperson durch die Intervention diesbezüglich verschlechtert. Die enthaltenen Teilbereiche des Reflektierens in den Definitionen nehmen bei neun Jugendlichen zu und bei 13 Testpersonen bleiben sie unverändert. Zwei Personen reduzieren die enthaltenen Teilbereiche von zwei auf einen. Bei den enthaltenen Teilbereichen gibt es überwiegend gleichbleibende Werte vor und nach der Intervention, gefolgt von Zunahmen der Teilbereiche.

Statistisch gesehen verändert sich die fachliche Nähe in den Definitionen der Schülerinnen und Schüler durch die Intervention signifikant positiv. Die Kontrollgruppe zeigt dabei keine statistisch messbaren Veränderungen. Hinsichtlich der enthaltenen Teilbereiche können weder bei der Experimentalgruppe noch bei der Kontrollgruppe signifikante Entwicklungen berichtet werden.

Die Teilforschungsfrage lässt sich anhand der qualitativen und quantitativen Ergebnisse wie folgt beantwortet: Das Begriffsverständnis von Reflektieren nimmt durch das Blogprojekt insgesamt zu. Die Zunahme zeigt sich hinsichtlich der fachlichen Nähe in den Definitionsformulierungen mit hohen Effektstärken von $d = 1,643$ bei den Testpersonen, die die Reflektiermethoden in der Reihenfolge Gibbs - Hüte erhalten und von $d = 1,454$ bei denjenigen, die sie in umgekehrter Reihenfolge erfahren. Die meisten Schülerinnen und Schüler können Reflektieren nach der Intervention präziser definieren als zuvor. Bei den enthaltenen Teilbereichen gibt es qualitativ gesehen eine gleichbleibende bis verbessernde Tendenz, die quantitativ jedoch nicht bestätigt werden kann.

(5) Lässt sich ein Zusammenhang zwischen der Reihenfolge des Erhaltens der Reflektiermethoden und der Veränderung der Reflektierfähigkeiten insgesamt feststellen?

Für die Beantwortung dieser Frage ist die quantitative Auswertung von Bedeutung. Die Ergebnisse der qualitativen Analyse zeigen, dass die Testpersonen positive Veränderungen der Reflektierfähigkeit durch die Intervention aufweisen. Dies wird durch die Untersuchung der einzelnen Kriterien der Scores, maximal erreichten und am häufigsten erreichten Stufen der Reflexion- und der Reflexivitätsfähigkeit ersichtlich. Ob bei der Entwicklung die Reihenfolge des Erlernens der beiden Reflektiermethoden von Bedeutung ist, wurde explorativ im Rahmen der quantitativen Auswertung untersucht. Folgendes Ergebnis bleibt festzuhalten: Die Reihenfolge Hütemethode - Gibbs Reflexivitätszyklus führt insgesamt zu einer stärkeren Verbesserung der Reflektierfähigkeit als die Reihenfolge Gibbs Reflexivitätszyklus - Hütemethode. Die Effekte der Intervention auf den Gesamtscore im Reflektieren, den Reflexionsscore, die maximal erreichte Reflexionsstufe, die maximal erreichte Reflexivitätsstufe, die am häufigsten erreichte Reflexionsstufe und die am häufigsten erreichte Reflexivitätsstufe sind bei der ersten Reihenfolge höher als bei der zweiten hinsichtlich des Vergleichs der Ergebnisse im ersten und letzten Beitrag. Der Gesamtscore im Reflektieren erfährt mit einem Effekt von $d = 4,772$ die deutlichste positive Beeinflussung durch die Intervention bei den Schülerinnen und Schülern, die zuerst die Hütemethode als Reflektierhilfe erhalten. Der gleiche Effekt beträgt bei den Teilnehmenden, die zuerst Gibbs Reflexivitätszyklus erhalten, $d = 4,061$.

Eine zusätzliche Erkenntnis liefert der Vergleich der durchschnittlichen Werte von Textblock 1 (vor der ersten Teilintervention) und Textblock 3 (nach beiden Teilinterventionen). Hier ist auch der größte Effekt auf den Gesamtscore im Reflektieren bei den Testpersonen, die die Reflektiermethoden in der Reihenfolge Hüte - Gibbs erlernen, feststellbar mit $d = 3,362$. Der entsprechende Effekt bei denjenigen, die sie in umgekehrter Reihenfolge erhalten, beträgt $d = 3,217$ und ist leicht niedriger, aber dennoch hoch. Auch die maximal erreichte Reflexivitätsstufe und die am häufigsten erreichte Reflexivitätsstufe verändert sich von Textblock 1 auf 3 bei der ersten Gruppe deutlich höher als bei der Gruppe, die Gibbs Reflexivitätszyklus zuerst erhalten.

Hinsichtlich der Verbesserung des Reflexionsscores und der maximal erreichten Reflexionsstufe schneidet die Gruppe, die die Hütemethode als erste Reflektierhilfe erhält, schlechter ab als die Gruppe, die Gibbs Reflexivitätszyklus zuerst erhält. Diese Ergebnisse aus dem Vergleich der Werte von Textblock 1 und 3 stimmen nicht mit den Ergebnissen des Vergleichs der Werte des ersten und letzten Beitrags überein. Ebenso ist die Entwicklung des Reflexivitätsscores zwischen den beiden Vergleichen unterschiedlich: Dieser verbessert sich bei der Gruppe, die die Hütemethode zuerst bekommt, von Textblock 1 auf 3 deutlich stärker als bei der anderen Gruppe.

Aus den beiden quantitativen Explorationen geht hervor, dass der Gesamtscore im Reflektieren bei den Probandinnen und Probanden, die die Hütemethode als erste und Gibbs Reflexivitätszyklus als zweite Reflektiermethode erlernen, immer einen höheren Effekt durch die Intervention aufweist als bei den Testpersonen, die die Reihenfolge in umgekehrter Reihenfolge bekommen. Auch die maximal erreichte Reflexivitätsstufe und die am häufigsten erreichte Reflexivitätsstufe entwickelt sich bei der Hüte-Gibbs Reihenfolge bei beiden Vergleichen stärker als bei der Gibbs-Hüte Reihenfolge. Die Reihenfolge ist folglich ein Faktor, der die Entwicklung der Reflektierfähigkeit beeinflusst. Die Untersuchung gibt Anlass dazu, folgende Hypothese aufzustellen, die in weiteren Studien noch konkreter untersucht werden muss: Die Reihenfolge Hütemethode - Gibbs Reflexivitätszyklus erzielt höhere Erfolge hinsichtlich der Reflektierfähigkeit als die Reihenfolge Gibbs Reflexivitätszyklus - Hütemethode. Die Förderung der Reflexionsfähigkeit und die anschließende Förderung der Reflexivitätsfähigkeit zur Förderung der gesamten Reflektierfähigkeit scheint nach den Erkenntnissen dieser Studie eine sinnvolle Strategie zu sein.

Die unterschiedlichen Ergebnisse der Vergleiche der Textblöcke und der einzelnen Beiträge, die im Unterricht zum selben Thema verfasst werden, lassen darauf schließen, dass die zuletzt erhaltene Reflektierhilfe bei den direkt darauffolgenden Beiträgen noch stärkeren Einfluss auf das Reflektieren hat als die Hilfe, die vorher gegeben wird. Darauf wird im nachfolgenden Kapitel noch genauer eingegangen (vgl. Kapitel 5.1).

(6) Unterscheidet sich das Potenzial beider Reflektiermethoden insgesamt hinsichtlich des Geschlechts? Wenn ja, wodurch?

Dieser Frage wurde im Rahmen der quantitativen Untersuchung durch die explorative Generierung einer neuen Hypothese nachgegangen. Dabei kommt man zum Ergebnis, dass sich das Potenzial der Gesamtintervention unterschiedlich auf die Geschlechter auswirkt. Während bei beiden eine deutliche Gesamtverbesserung des Gesamtscores im Reflektieren besteht mit Effektstärken von $d = 4,353$ der männlichen und $d = 3,952$ der weiblichen Testpersonen, verändern sich die beiden Teilbereiche Reflexion und Reflexivität unterschiedlich. Die Reflexivitätsfähigkeit gemessen am Reflexivitätsscore und der am häufigsten erreichten Reflexivitätsstufe verändert sich nur bei den weiblichen Testpersonen signifikant mit Effektstärken von

$d = 1,119$ und $d = 1,692$. Auf die Variable maximal erreichte Reflexivitätsstufe ist ebenfalls bei den weiblichen Testpersonen ein höherer Effekt der Gesamtintervention als bei den männlichen feststellbar. Hinsichtlich der Reflexionsfähigkeit verhält es sich genau umgekehrt: Hier verbessert sich der Reflexionsscore und die maximal erreichte Reflexionsstufe der männlichen gegenüber den weiblichen Teilnehmenden deutlich mehr mit Effektstärken von $d = 2,071$ und $d = 1,993$. Die am häufigsten erreichte Reflexionsstufe nimmt nur bei den männlichen Personen signifikant zu. Die Teilforschungsfrage wird deshalb mit ja beantwortet. Eine neue Erkenntnis ist, dass sich weibliche Testpersonen hinsichtlich des Nachdenkens über eigene Handlungen und männliche Teilnehmende hinsichtlich des Nachdenkens über Themen stärker verbessern.

(7) Besteht ein Zusammenhang zwischen der eigenen Bewertung der Reflektiermethoden und der Veränderung der Reflektierfähigkeit durch diese?

Auf diese Frage kann im Rahmen dieser Studie keine Antwort gegeben werden, da die beiden Stichproben, die sich aus den Angaben des Fragebogens am Ende des Blogprojekts ergeben, zu unterschiedlich ausfallen. Eine valide, statistische Aussage kann deshalb nicht generiert werden.

(8) Inwieweit verändern sich persönliche Einstellungen gegenüber dem Reflektieren durch das Blogprojekt insgesamt?

Für die Beantwortung dieser Frage wurden die im Fragebogen erhobenen, persönlichen Einschätzungen zu folgenden Aspekten des Reflektierens genauer betrachtet:

- Ich treffe Entscheidungen eher spontan, ohne viel Nachzudenken.
- Ich habe immer eine Erklärung für mein Verhalten.
- Ich denke oft über mein Verhalten nach, um die Gründe dafür verstehen zu können.
- Ich akzeptiere andere Ansichten oder Meinungen bei komplexen Themen.
- Manchmal frage ich mich, wie andere etwas tun und versuche eine bessere Lösung zu finden.
- Ich denke oft in meiner Freizeit über Unterrichtsinhalte nach.

Zwischen beiden Experimentalgruppen und der Kontrollgruppe gibt es keine signifikanten Unterschiede bezüglich der persönlichen Einstellungen zum Reflektieren vor und nach der Intervention. Die Intervention nimmt folglich innerhalb dieser Studie keinen Einfluss auf wiedergegebene, persönliche Einstellungen zum Reflektieren im Fragebogen. Deshalb wird die Frage folgendermaßen beantwortet: Die persönlichen Einstellungen gegenüber dem Reflektieren verändern sich nicht durch das Blogprojekt.

Welche Potenziale bieten ausgewählte Reflektiermethoden für die gezielte Förderung von Reflektierfähigkeiten bestehend aus Reflexions- und Reflexivitätsfähigkeiten von Jugendlichen beim Bloggen im Rahmen einer Bildung für nachhaltige Entwicklung?

Die Potenziale von de Bonos Hütemethode und Gibbs Reflexivitätszyklus für die Förderung von Reflektierfähigkeiten von Schülerinnen und Schülern im Rahmen eines Blogprojekts innerhalb eines P-Seminars mit dem Schwerpunkt Nachhaltigkeit sind nach den Erkenntnissen dieser Studie sehr hoch, wie die Ergebnisse der Teilfragen zeigen. Während de Bonos Hütemethode einen großen Einfluss auf die Verbesserung der Reflexionsfähigkeiten hat, sorgt Gibbs Reflexivitätszyklus für eine Verbesserung der Reflexivitätsfähigkeiten. Insgesamt führen die beiden Reflektiermethoden zu einer deutlichen Steigerung der Reflektierfähigkeiten. Die Jugendlichen reflektieren häufiger auf höheren Reflektierniveaus, die sie teilweise vor der Intervention nicht erreicht haben. Die Potenziale beider Reflektierhilfen sind je nach Testperson unterschiedlich stark, insgesamt betrachtet überwiegen jedoch die deutlichen positiven Einflüsse beider auf die gezeigten Reflektierfähigkeiten. Auch das Verständnis von Reflektieren wird durch die Intervention insgesamt erweitert. Im nächsten Kapitel wird auf die Ergebnisse dieser Studie noch vertiefter eingegangen.

5 Diskussion

Die Studie hat sich das Ziel gesetzt, zu überprüfen, welche Potenziale die beiden Methoden Gibbs Reflexivitätszyklus und de Bonos Hütemethode auf die Reflektierfähigkeiten von Schülerinnen und Schüler haben. Die Erkenntnis ist eindeutig, dass sich beide Methoden als sinnvolle Instrumente zur Förderung dieser Fähigkeit eignen. Dies stimmt mit den Ergebnissen der bisher wenig vorhandenen Forschungsarbeiten über die Wirkung beider Reflektiermethoden überein (vgl. Kapitel 2.7.2.1). Auch die Häufigkeit des Einsatzes von de Bonos Hütemethode zur Förderung des kritischen Denkens und der Einsatz von Gibbs Reflexivitätszyklus zur Ausbildung eines *reflective practitioners* nach Schön (1983) kann mit den Ergebnissen dieser Studie nachvollzogen werden. Während Gibbs Reflexivitätszyklus eindeutig mehr Einfluss auf die Förderung der Reflexivität hat, steigert das Anwenden der Hütemethode mehr die Reflexion als Teilbereich des Reflektierens. Somit bestätigen sich die ersten Vermutungen der Wirkung dieser Methoden, die aus der Literaturrecherche und der bisherigen Verwendung dieser Methoden entwickelt wurden. Im Folgenden werden die Erkenntnisse dieser Studie noch genauer diskutiert.

5.1 Erkenntnisse aus der Studie

Neben der Erkenntnis, dass die Kombination beider Reflektiermethoden über einen längeren Zeitraum hinweg den gewünschten Effekt der allgemeinen Reflektierfähigkeitsförderung hat, können weitere Erkenntnisse gewonnen werden. Zum einen wird der Vorschlag abgeleitet, dass der Einsatz der Reflektiermethoden zur Förderung von Reflektierfähigkeiten kombiniert stattfinden sollte und nicht beide Methoden nacheinander erlernt werden sollten. Grund dafür sind die beobachteten stärker ausgeprägten Reflektieranteile innerhalb der Blogbeiträge direkt nach den Methoden, die beabsichtigen, diese auch zu fördern. Während dem anderen Teilbereich des Reflektierens bei Fokussierung auf eine Methode weniger Beachtung geschenkt wird, verbessert sich die Leistung des Reflektierbereichs, der durch die Methode auch gefördert werden soll, deutlich. Gibbs Reflexivitätszyklus führt folglich zu einer Förderung der Reflexivität, doch gleichzeitig nimmt die Reflexionsfähigkeit ab oder bleibt auf einem ähnlichen Niveau. Bei der Hütemethode ist es genau umgekehrt: Sie bewirkt generell eine Förderung der Reflexion, während sich die Reflexivitätsfähigkeit verschlechtert oder ähnlich bleibt. Als möglicher Grund dafür könnte die explizite Differenzierung zwischen dem Nachdenken über das Thema und dem Nachdenken über die Handlung innerhalb der Definition von Reflektieren und der Aufgabenformulierung sein. Dadurch, dass stets zwei Teilbereiche abgedeckt werden sollen, kann es zu stärkeren Fokussierungen auf nur einen Teilbereich kommen, besonders, wenn Methoden auch gezielt nur einen davon speziell fördern. Dies ist nach Erkenntnissen dieser Studie vor allem häufig der Fall, wenn kurz vor dem schriftlichen Reflektieren eine Reflektiermethode erlernt wurde. Die Schülerinnen und Schüler beschränken sich dann teilweise

auf nur diesen einen Teilbereich des Reflektierens und beachten den anderen Teilbereich nur weniger oder überhaupt nicht. Würde man die Förderung beider Teilbereiche in eine Reflektierhilfe aufnehmen, käme es vermutlich direkt zum Ausbau beider Reflektierarten, der Reflexion und der Reflexivität, da keine so explizite Differenzierung der beiden Teilbereiche vorgenommen werden würde. Eine solche Kombination könnte gut aus Gibbs Reflexivitätszyklus heraus entwickelt werden: Die Phase der Evaluation könnte intensiver aufgebaut sein und sich am Schema von de Bonos Hütemethode orientieren, um die Reflexion innerhalb der Reflexivität zu berücksichtigen.

Reinmann (2005) und Smyth (1992) sind der Meinung, dass Reflektieren unterschiedlich aufgefasst werden kann, je nachdem, worauf man sich dafür konkret bezieht. Auch bei den Schülerinnen und Schülern im Rahmen dieser Studie wird ersichtlich, dass unterschiedliche Verständnisse des Begriffs vorhanden sind. Manche beziehen sich dabei auf den physikalischen Begriff, andere versuchen, ihn auf kognitive Vorgänge zu übertragen. Das Ergebnis dabei ist ernüchternd: Vor der Intervention zeigen die Jugendlichen deutliche Schwächen beim Verständnis vom Reflektieren und verstehen darunter häufig ein reines Wiedergeben von Inhalten. Obwohl es sich um Jugendliche der Oberstufe handelt, die kurz vor Abschluss ihrer Schullaufbahn stehen, ist vielen von ihnen nicht klar, was Reflektieren eigentlich bedeutet. Dies deckt sich mit der vielfältigen Verwendung des Begriffs je nach Fachdisziplin und der oft ungenauen Begriffsklärung (vgl. Abou Baker El-Dib 2007). Als Erkenntnis dieser Beobachtung kann schlussgefolgert werden, dass die Fähigkeit, adäquat über Themen und Handlungen reflektieren zu können, bisher nicht ausreichend im schulischen Kontext gefördert wird. Natürlich kann hierbei nicht von einer allgemeinen Feststellung gesprochen werden, sondern von der Erkenntnis der drei Schulklassen unterschiedlicher Gymnasien. Obwohl Reflektieren als eine Grundfähigkeit für den Kompetenzerwerb gesehen werden kann (vgl. Rychen 2003) und im Bereich der BNE und des Geographieunterrichts von Bedeutung ist, wurde sie scheinbar in allen drei Klassen nie explizit trainiert. Mit dieser Arbeit sollen Lehrende dafür sensibilisiert werden, der Fähigkeit des Reflektierens mehr Aufmerksamkeit zu widmen und diese häufiger explizit im Unterricht aufzunehmen. Die beiden Reflektiermethoden, die im Rahmen dieser Arbeit entwickelt wurden, können dafür als Methodik verwendet werden. Die für diese Studie verwendeten Arbeitsblätter wurden speziell für das Reflektieren im Rahmen einer BNE entwickelt, sie können aber mit minimaler Überarbeitung (Streichung der Stellen, die explizit Bezug auf das Konzept des Nachhaltigkeitsdreiecks nehmen) auch für die Förderung von Reflektierfähigkeiten in anderen Kontexten verwendet werden.

Interessant ist weiterhin die Erkenntnis, dass sich die Reflektierfähigkeiten weiblicher und männlicher Testpersonen durch die Intervention unterschiedlich entwickeln. Auch die unterschiedlichen Entwicklungen der einzelnen Probandinnen und Probanden, bei denen einige sehr große Verbesserungen zeigen und manche weniger große, betonen die Wichtigkeit, dass

Methoden nicht bei allen die gleiche Wirkung erzielen. Die Heterogenität von Schülerinnen und Schülern sollte deshalb, wie es in der Literatur häufig empfohlen wird (vgl. Vock und Gronostaj 2017, S. 63 ff.), stets berücksichtigt und unterschiedliche Lernarrangements für die individuelle Lernförderung angeboten werden. Dabei gilt ebenfalls zu beachten, dass geschlossene Reflexionen, wie Wyss (2008, S. 6) Reflektierprozesse, angeleitet durch Strukturierungshilfen, nennt, den Vorgang des Reflektierens unterstützen aber gleichzeitig auch als Einschränkungen der eigenen Denkkideen eingestuft werden können. Um der Diversität der Schülerinnen und Schüler gerecht zu werden, sollten die Reflektiermethoden deshalb stets als optionale Angebote für die Jugendlichen verstanden werden. Man sollte sie nicht durch beispielsweise den Druck durch Notengebung dazu nötigen, diese zwingend anzuwenden, wenn sie für sie nicht als Hilfe angesehen werden.

Keiner der Teilnehmenden hat im Laufe des Projekts die höchstmögliche Stufe beim Reflektieren über Handlungen erreicht. Die Professionalisierung scheint deshalb erst erreichbar zu sein, wenn längerfristig über das gleiche Verhalten reflektiert wird, bis dieses schließlich professionalisiert werden kann. Da die Schülerinnen und Schüler in diesem Projekt stets über neue Themen und ihre Handlungen diesbezüglich reflektiert haben, konnte sich auch mithilfe der Reflektierhilfen keine Professionalisierung ausbilden. Hinsichtlich der Reflexion wurde nach dem Erlernen der Reflektiermethoden die höchste Stufe mehrmals erreicht. Dies deutet darauf hin, dass die Reflexion leichter zu erlernen ist als die Reflexivität. Eine andere Erklärung für diese Feststellung könnte auch darin begründet sein, dass die Probandinnen und Probanden insgesamt bevorzugt allgemein über Themen reflektieren und sich selbständig Gedanken über ihr Verhalten machen, dies aber nur ungern zu Papier bringen.

Eine weitere Erkenntnis ist, dass die Methoden und das Blogprojekt insgesamt keinen Einfluss auf die persönlichen Einschätzungen zum Reflektieren nehmen. Die Schülerinnen und Schüler geben vor und nach dem Projekt ähnliche Einschätzungen zum Vorgehen beim Reflektieren und der Häufigkeit von Reflektiervorgängen ab. Die langfristige, reflexive Auseinandersetzung mit unterschiedlichen Themen scheint folglich keinen Einfluss auf eigene Einstellungen und Verhaltensweisen bezüglich des Reflektierens zu nehmen. Zwar ändert sich die Qualität der Reflektierfähigkeit durch die Methoden, eine Zunahme an Reflektiervorgängen außerhalb des Projekts gibt es jedoch scheinbar nicht. Die Ergebnisse beruhen wie bereits erwähnt auf den Selbsteinschätzungen der Teilnehmenden. Eine Änderung der tatsächlichen Einstellung und des tatsächlichen Verhaltens diesbezüglich kann jedoch trotzdem unbewusst stattgefunden haben.

Die Durchführung des Blogprojekts insgesamt hat wie erwartet funktioniert und wurde von den Schülerinnen und Schülern positiv angenommen. Die Stunden, in denen die Reflektierhilfen behandelt wurden, verliefen ohne Probleme und es gab keine Rückfragen zu den Methoden.

Ein Blogprojekt in Kombination mit expliziten Methoden zur Förderung von Reflektierfähigkeiten wird folglich als geeignete Form für das Vertiefen von Reflektierprozessen gesehen. Auch der Einsatz von Blogs als Darstellungsoberfläche für Reflektiervorgänge (vgl. Kapitel 2.5.1) hat sich im Rahmen der Studie als sinnvoll herausgestellt. Die beiden Modelle, die in Rahmen der Arbeit verwendet wurden, werden ebenfalls als geeignet eingestuft, um gezeigte Reflektierfähigkeiten messen zu können. Wichtig ist dabei zu beachten, dass die Einstufung der Reflektierfähigkeiten auf schriftlichen Arbeiten der Schülerinnen und Schüler beruht und diese von der individuellen Ausdrucksfähigkeit und Performanz abhängen. Ob die Texte auch die tatsächlichen kognitiven Fähigkeiten abbilden, bleibt unklar.

Während des Blogprojekts wurde deutlich, dass vielen Probandinnen und Probanden nicht bewusst war, was unter dem Konzept des Nachhaltigkeitsdreiecks verstanden wird. Obwohl vorab mit den Lehrkräften besprochen wurde, das Konzept vor Beginn des Projekts ausführlich mit den Schülerinnen und Schülern zu behandeln, ist dies scheinbar nicht in allen Klassen erfolgt. Als Erkenntnis kann deshalb festgehalten werden, dass zu erwartende Grundvoraussetzungen bei Forschungsarbeiten am besten selbst überprüft und bearbeitet oder für die Lehrkräfte eine fertige Unterrichtsstunde dafür entwickelt werden sollte. Somit kann sichergestellt werden, dass die vordefinierten Voraussetzungen für eine Studie auch tatsächlich geschaffen werden und dies nicht nur theoretisch erfolgt.

5.2 Grenzen der Studie

Diese Studie konzentriert sich darauf, die Potenziale zweier Reflektiermethoden auf die Reflektierfähigkeiten bestehend aus den Teilbereichen der Reflexions- und Reflexivitätsfähigkeit zu untersuchen. Deshalb wurden dabei überwiegend die Veränderungen der entsprechenden Fähigkeiten (gemessen in Scores und maximal erreichte und am häufigsten erreichte Stufen) im Sinne eines Prä-Post-Test-Designs, betrachtet. Die vorhandenen Reflektierfähigkeiten der Testpersonen zu den unterschiedlichen Messzeitpunkten wurden zwar im Rahmen der Auswertung beachtet, allerdings nicht selbst weiter analysiert, wie es beispielsweise Brendel (2017) tut. Ein Ansatz für weiterführende Untersuchungen wäre die Klärung der Frage, inwieweit Entwicklungen von Reflektierfähigkeiten durch das explizite Fördern dieser von Ausgangswerten abhängen. Im Rahmen dieser Studie ist die Stichprobe für quantitative Aussagen diesbezüglich zu klein, die qualitativen Daten könnten dafür jedoch herangezogen werden.

Weiterhin bleibt im Rahmen dieser Forschung unklar, inwieweit das regelmäßige Feedback in Form von Kommentaren zusätzlich Einfluss auf die Förderung der Reflektierfähigkeiten nimmt. Für Feedback wurde sich bewusst entschieden, da nach Abels (2011, S. 68) die soziale Komponente beim reflexiven Schreiben als wichtig erachtet wird. Der virtuelle Dialog, der dadurch entsteht, sollte die Motivation während des Projekts aufrechterhalten. Die Rückmeldungen beschränkten sich zwar stets auf das Anwenden der erlernten Reflektiermethoden und an das

Erinnern beider Teilbereiche des Reflektierens abgeleitet von der Definition, die mit den Teilnehmenden besprochen wurde, sodass sie keinen Mehrwert außerhalb der Reflektierhilfen liefern sollen. Ob dies jedoch tatsächlich so ist, konnte in dieser Studie nicht untersucht werden. Ein Vergleich zweier Gruppen, wovon nur eine Gruppe Feedback während des Anwendens von Reflektiermethoden erhält, könnte dieser Überlegung Antworten liefern und stellt somit ein weiterzuverfolgendes Forschungsgebiet dar.

Für die Auswertung wurde sich an der Kombination qualitativer und quantitativer Forschungstraditionen orientiert, um eine möglichst ganzheitliche, multiperspektivische Perspektive auf den Untersuchungsgegenstand einnehmen zu können. Durch die Kombination musste die Stichprobe so gewählt werden, dass sie im qualitativen Auswertungsrahmen von der Größe her noch auswertbar und gleichzeitig für eine statistische Auswertung groß genug ist. Dabei sei angemerkt, dass die Ergebnisse dieser Studie nicht als verallgemeinerbar angesehen werden, sondern sich nur auf Erkenntnisse innerhalb dieses Rahmens beziehen. Die entwickelten Ergebnisse sollen vielmehr Anlass dafür geben, die wichtigen Themenfelder des Reflektierens und der Förderung von Reflektierfähigkeiten weiter zu erforschen. Die Erkenntnisse, dass sowohl Reflexions- als auch Reflexivitätsfähigkeiten gefördert werden können und sich Gibbs Reflexivitätszyklus und die Hütemethode hierfür als sinnvolle Instrumente herausgestellt haben, sollen eine der noch vielen bestehenden Lücken innerhalb des Forschungsgebiets des Reflektierens im schulischen Kontext schließen und für weitere Forschungsarbeiten diesbezüglich anschlussfähig sein.

Das Reflektieren im Rahmen dieser Studie wurde innerhalb einer BNE untersucht und fokussiert sich auf das tiefgründige Nachdenken über Themen und Handlungen mit Berücksichtigung des Nachhaltigkeitskonzepts. Hierbei haben sich die Reflektiermethoden bewährt und zeigen hohe Potenziale in der Förderung des Reflektierens. Ob dies auch in anderen schulischen Kontexten, wie beispielsweise dem interkulturellen Lernen oder bei der Durchführung eigener Experimente in den Naturwissenschaften der Fall ist, kann ebenfalls in weiteren Studien erforscht werden.

6 Fazit

Um der Bedeutung der Fähigkeit des Reflektierens im schulischen und auch außerschulischen Bereich mehr Stimme zu verleihen, wurde für diese Dissertationsarbeit der Untersuchungsgegenstand Förderung von Reflektierfähigkeiten gewählt. Die Studie ging der Frage nach, welche Potenziale Gibbs Reflexivitätszyklus und de Bonos Hütemethode für die Förderung der Reflektierfähigkeiten bestehend aus den Fähigkeiten der Reflexion und der Reflexivität, bieten. Die Erkenntnis ist eindeutig: Die beiden Reflektiermethoden haben sich als geeignete Instrumente herausgestellt, um das Reflektieren mit Schülerinnen und Schülern innerhalb einer BNE zu fördern. Gibbs Reflexivitätszyklus bewirkt vor allem eine Steigerung der Reflexivitätsfähigkeit, während die Hütemethode besonders die Reflexionsfähigkeit verbessert. Folglich eignen sich beide gewählten Methoden für die Verwendung im Unterricht und können effizient und unkompliziert eingesetzt werden. Die beiden Niveaustufenmodelle, die im Rahmen der Auswertung zur Bestimmung der Reflexions- und Reflexivitätsfähigkeiten gewählt wurden, bieten ein von der Theorie heraus entwickeltes, valides Konstrukt für Evaluationen von Reflektierfähigkeiten. Abschließend sei daran erinnert, dass das Reflektieren eine Fähigkeit ist, die für lebenslanges Lernen und den Kompetenzerwerb unerlässlich ist. Sie sollte deshalb nicht nur als Mitläuferfähigkeit im Unterricht gesehen werden, sondern ihr sollte mehr Aufmerksamkeit geschenkt und sie als eigene Kompetenz eingeübt werden, ähnlich wie bei der Arbeit mit Klimadiagrammen: Es ist unmöglich ein Klimadiagramm richtig lesen und auswerten zu können, ohne zu wissen, aus welchen Bestandteilen ein Diagramm im Allgemeinen besteht und wie es richtig gelesen wird. Ohne das Grundverständnis und -werkzeug wird man zu keinen hochwertigen Ergebnissen kommen. Wie soll schließlich ein Kompetenzerwerb erfolgreich gelingen, wenn dafür eine der Grundfähigkeiten, gemeint ist hierbei die Reflektierfähigkeit, nicht geklärt und eingeübt wird?

Literaturverzeichnis

- Abels, S. (2011). *LehrerInnen als "Reflective Practitioner"*. Verlag für Sozialwissenschaften. Wiesbaden.
- Abou Baker El-Dib, M. (2007). Levels of reflection in action research. An overview and an assessment tool. *Teaching and Teacher Education*, S. 24–35.
- Akreimi, L. (2019). Stichprobenziehung in der qualitativen Sozialforschung. In N. Baur, & J. Blasius (Hrsg.), *Handbuch Methoden der empirischen Sozialforschung*. Wiesbaden: Springer Fachmedien Wiesbaden, S. 313–331.
- Backhaus, K., Erichson, B., Plinke, W., & Weiber, R. (2000). *Multivariate Analysemethoden: Eine anwendungsorientierte Einführung*. Berlin, Heidelberg: Springer.
- Bade, P., Herold, G., & Lehmbäcker, E. (2018). Reflexionskompetenz fördern: Reflexion und Reflexionskompetenz in der Lehrerbildung. Landesinstitut für Lehrerbildung und Schulentwicklung. <http://edoc.sub.uni-hamburg.de/hlb/volltexte/2018/198/>. Gesehen 5. Dezember 2019.
- Bahr, M. (2013). Bildung für nachhaltige Entwicklung (BNE). In M. Rolfes, & A. Uhlenwinkel (Hrsg.), *Metzler Handbuch 2.0 Geographieunterricht: Ein Leitfaden für Praxis und Ausbildung*. Braunschweig: Westermann, S. 17–23.
- Bain, J., Ballantyne, R., Mills, C., & Lester, N. (2002). *Reflecting on Practice: Student Teachers' Perspectives*. Flaxton: Post Pressed.
- Bain, J., Ballantyne, R., Packer, J., & Mills, C. (1999). Using Journal Writing to Enhance Student Teachers' Reflectivity During Field Experience Placements. *Teachers and Teaching*, S. 51–73.
- Bamberg, G., Baur, F., & Krapp, M. (2017). *Statistik: Eine Einführung für Wirtschafts- und Sozialwissenschaftler*. Berlin/Boston: De Gruyter.
- Barth, C. (2018). Selbstreflexion. In C. Barth (Hrsg.), *Kompetenzentwicklung im Studium ermöglichen*. Wiesbaden: Springer Fachmedien Wiesbaden, S. 27–34.
- Barth, M. (2016). Forschung in der Bildung für nachhaltige Entwicklung: Entstehung und Verortung eines Forschungszweiges. In M. Rieckmann, & M. Barth (Hrsg.), *Empirische Forschung zur Bildung für nachhaltige Entwicklung - Themen, Methoden und Trends*. Opladen: Verlag Barbara Budrich, S. 37–50.
- Barth, M., & Rieckmann, M. (2016). Einleitung: Eine Summer School zur empirischen Forschung zur Bildung für nachhaltige Entwicklung. In M. Rieckmann, & M. Barth (Hrsg.),

- Empirische Forschung zur Bildung für nachhaltige Entwicklung - Themen, Methoden und Trends.* Opladen: Verlag Barbara Budrich, S. 9–14.
- Bergmann, R. (2009). *Medienkompetenz: Digitale Medien in Theorie und Praxis für sozialpädagogische Berufe.* Troisdorf: Bildungsverlag EINS.
- Blanca, M. J., Alarcón, R., Arnau, J., Bono, R., & Bendayan, R. (2017). Non-normal data: Is ANOVA still a valid option? *Psicothema*, S. 552–557.
- Bloom, B. S. (1956). *Taxonomy of educational objectives: The classification of educational goals.* Benjamin S[amuel] Bloom, ed. New York usw.: Longmans Green & Co.
- Bock, I. (2001). Pädagogische Anthropologie. In L. Roth (Hrsg.), *Pädagogik. Handbuch für Studium und Praxis.* München: Oldenburg, S. 112–122.
- Bono, E. de (1992a). *Six Thinking Hats for Schools: Book 2.* Cheltenham: Hawker Brownlow Education.
- Bono, E. de (1992b). *Six Thinking Hats for Schools: Book 3.* Cheltenham: Hawker Brownlow Education.
- Bono, E. de (1992c). *Six Thinking Hats for Schools: Book 4.* Cheltenham: Hawker Brownlow Education.
- Bono, E. de (2000). *Six thinking hats.* London: Penguin.
- Bono, E. de (2005). *De Bonos neue Denkschule: Kreativer denken, effektiver arbeiten, mehr erreichen.* Frankfurt am Main: mvgVerl.
- Bono, E. de (2017). *Six thinking hats.*
- Borton, T. (1970). *Reach, touch, and teach: Student concerns and process education.* New York: McGraw-Hill Paperbacks.
- Bortz, J., & Döring, N. (2006). *Forschungsmethoden und Evaluation.* Berlin, Heidelberg: Springer Berlin Heidelberg.
- Bortz, J., & Schuster, C. (2010). *Statistik für Human- und Sozialwissenschaftler.* Berlin, Heidelberg: Springer-Verlag Berlin Heidelberg.
- Boud, D., Cohen, R., & Walker, D. (1993). *Using experience for learning.* Buckingham: Open University Press.
- Bräuer, G. (2014). *Das Portfolio als Reflexionsmedium für Lehrende und Studierende.* Opladen, Toronto: Budrich.

- Brendel, N. (2017). *Reflexives Denken im Geographieunterricht: Eine empirische Studie zur Bestimmung von Schülerreflexion mithilfe von Weblogs im Kontext Globalen Lernens*. Münster: Waxmann Verlag GmbH.
- Brendel, N., & Schrüfer, G. (2013). Web 2.0. In D. Böhn, & G. Obermaier (Hrsg.), *Wörterbuch der Geographiedidaktik: Begriffe von A - Z*. Braunschweig: Westermann, S. 289–290.
- Brendel, N., & Schrüfer, G. (2018). Weblogs als Reflexionsmedium im Globalen Lernen. *Globales Lernen im digitalen Zeitalter*, S. 79–106.
- Brosius, H.-B., & Koschel, F. (2012). Messen und Zählen I. In H.-B. Brosius, & F. Koschel (Hrsg.), *Methoden der empirischen Kommunikationsforschung*. Wiesbaden: VS Verlag für Sozialwissenschaften, S. 51–63.
- Brüsemeister, T. (2008). *Qualitative Forschung: Ein Überblick*. Wiesbaden: VS Verlag für Sozialwissenschaften.
- Buddeberg, M. (2014). *Zur Implementation des Konzepts Bildung für nachhaltige Entwicklung: Eine Studie an weiterführenden Schulen in Nordrhein-Westfalen*. Münster: Waxmann.
- Bühl, A. (2012). *SPSS 20: Einführung in die moderne Datenanalyse*. München: Pearson.
- Bühner, M., & Ziegler, M. (2009). *Statistik für Psychologen und Sozialwissenschaftler*. München: Pearson Deutschland; Pearson Studium.
- Bünig, H. (1991). *Robuste und adaptive Tests*. Berlin: De Gruyter.
- Carrington, S., & Selva, G. (2010). Critical Social Theory and Transformative Learning: Evidence in Pre-Service Teachers' Service-Learning Reflection Logs. *Higher Education Research and Development* 29, S. 45–57.
- Cendon, E., & Bischoff, F. (2014). Reflexives Lernen. Berufliche Praxis reflexiv betrachtet. In E. Cendon, & L. B. Flacke (Hrsg.), *Lernwege gestalten: Studienformate an der Schnittstelle von Theorie und Praxis. Veranstaltung der wissenschaftlichen Begleitung, 5. bis 6. Dezember 2013, Berlin*, S. 31–42.
- Chen, N.-S., Wei, C.-W., Wu, K.-T., & Uden, L. (2009). Effects of high level prompts and peer assessment on online learners' reflection levels. *Computers & Education* 52, S. 283–291.
- Chong, M. (2009). Is Reflective Practice a Useful Task for Student Nurses? *Asian Nursing Research*, S. 111–120.
- Cleff, T. (2019). *Angewandte Induktive Statistik und Statistische Testverfahren*. Wiesbaden: Springer Fachmedien Wiesbaden.

- Cohen, J. (1988). *Statistical power analysis for the behavioral sciences*. Hillsdale, N.J.: L. Erlbaum Associates.
- Copeland, W. D., Birmingham, C., La Cruz, E. D., & Lewin, B. (1993). The reflective practitioner in teaching: Toward a research agenda. *Teaching and Teacher Education* 9, S. 347–359.
- Creswell, J. W., & Plano Clark, V. L. (2011). *Designing and conducting mixed methods research*. Los Angeles, London, New Dehli, Singapore, Washington DC: Sage.
- Dewey, J. (1910). *How we think*. Boston: D.C. Heath.
- Dewey, J. (1933). *How we think: A restatement of the relation of reflective thinking to the educative process*. Boston: D.C. Heath and company.
- Dewey, J. (1938). *Experience and Education*. New York: Macmillan.
- Dewey, J. (2009). *Democracy and Education*. Portland: The Floating Press.
- DGfG (2014). Bildungsstandards im Fach Geographie für den Mittleren Schulabschluss. http://dgfg.geography-in-germany.de/wp-content/uploads/geographie_bildungsstandards.pdf. Gesehen 5. Dezember 2019.
- Döring, N., & Bortz, J. (2016). *Forschungsmethoden und Evaluation in den Sozial- und Humanwissenschaften*. Berlin, Heidelberg: Springer Berlin Heidelberg.
- Dubs, R. (1992). Die Förderung des kritischen Denkens in Unterricht. *Bildungsforschung und Bildungspraxis* 14, S. 28–56.
- Eickelmann, B., & Schulz-Zander, R. (2010). *Bildung und Schule auf dem Weg in die Wissensgesellschaft: [Festschrift für Prof. 'in Dr. Renate Schulz-Zander zum 65. Geburtstag]*.
- Eldeen, A. S. M., & Maher, A. E. (2016). The effect of the six thinking hats strategy in teaching health and fitness course on the development of creative thinking and the academic achievement level. *Science, Movement and Health* 16, S. 209–215.
- Eysel, C. (2006). *Interdisziplinäres Lehren und Lernen in der Lehrerbildung*. Zugl.: Heidelberg, Pädag. Hochsch., Diss., 2006. Logos-Verl. Berlin.
- Facione, P. A. (1990). *Critical Thinking: A Statement of Expert Consensus for Purposes of Educational Assessment and Instruction*. Newark: American Philosophical Association.
- Fakude, L., & Bruce, J. (2003). Journaling: A quasi-experimental study of student nurses' reflective learning ability. *Curationis* 26, S. 49–55.
- Faßler, M. (1997). *Was ist Kommunikation?* München: Fink.

- Feindt, A. (2007). *Studentische Forschung im Lehramtsstudium: Eine fallrekonstruktive Untersuchung studienbiografischer Verläufe und studentischer Forschungspraxen*. Leverkusen: Verlag Barbara Budrich.
- Felten, R. v. (2005). *Lernen im reflexiven Praktikum: Eine vergleichende Untersuchung*. Zugl.: Bern, Univ., Diss., 2004. Münster: Waxmann.
- Field, A. (2013). *Discovering statistics using SPSS: (and sex and drugs and rock 'n' roll)*. Los Angeles: Sage Publ.
- Finch, H. (2005). Comparison of the Performance of Nonparametric and Parametric MANOVA Test Statistics when Assumptions Are Violated. *Methodology: European Journal of Research Methods for the Behavioral and Social Sciences*, S. 27–38.
- Fishbein, M., & Ajzen, I. (1975). *Belief, attitude, intention and behavior: An introduction to theory and research*. Reading, Mass.: Addison-Wesley.
- Flick, U. (1996). *Qualitative Forschung: Theorie, Methoden, Anwendung in Psychologie und Sozialwissenschaften*. Reinbek bei Hamburg: Rowohlt-Taschenbuch-Verl.
- Flick, U. (2010). Design and process in qualitative research. In U. Flick, E. v. Kardorff, & I. Steinke (Hrsg.), *A companion to qualitative research*. London: Sage.
- Flick, U. (2017). Design und Prozess qualitativer Forschung. In , *Qualitative Forschung : ein Handbuch*. Reinbek bei Hamburg: rowohlts enzyklopädie im Rowohlt Taschenbuch Verlag, S. 252–265.
- Früh, W. (2007). *Inhaltsanalyse: Theorie und Praxis*. Konstanz: UVK-Verl.-Ges.
- Gibbs, G. (1988). *Learning by doing: A guide to teaching and learning methods*. Oxford: Further Educational Unit.
- Gibbs, G. (2013). Learning by doing. Oxford Brookes University. <https://thoughtsmostlyaboutlearning.files.wordpress.com/2015/12/learning-by-doing-graham-gibbs.pdf>.
- Girwidz, R. (2002). Neues Lernen mit neuen Medien? *Naturwissenschaften im Unterricht Physik 13*, S. 4–6.
- Glaser, B., & Strauss, A. L. (1967). *Discovery of Grounded Theory: Strategies for Qualitative Research*. Somerset: Taylor and Francis.
- Gläser-Zikuda, M., & Hascher, T. (2007). Zum Potenzial von Lerntagebuch und Portfolio. In M. Gläser-Zikuda, & T. Hascher (Hrsg.), *Lernprozesse dokumentieren, reflektieren und beurteilen. Lerntagebuch & Portfolio in Bildungsforschung und Bildungspraxis*. Bad Heilbrunn: Klinkhardt, S. 9–21.

- Glesner Fines, B. (2014). Assessing Reflection. <https://www.smu.edu/-/media/Site/Law/faculty/teaching-resources/Student-Reflection-Rubric.pdf>. Gesehen 5. Dezember 2019.
- Grobbauer, H., & Thaler, K. (2010). Globales Lernen - die Welt deuten, erfahren und verstehen. In G. Schröder, & I. Schwarz (Hrsg.), *Globales Lernen: Ein geographischer Diskursbeitrag*. Münster, New York, München, Berlin: Waxmann, S. 125–147.
- Gryl, I. (2012). Geographielehrende, Reflexivität und Geomedien. Zur Konstruktion einer empirisch begründeten Typologie. *Geographie und Ihre Didaktik* 40, S. 161–182.
- Haan, G. de (2008). Gestaltungskompetenz als Kompetenzkonzept der Bildung für nachhaltige Entwicklung. In , *Kompetenzen der Bildung für nachhaltige Entwicklung : Operationalisierung, Messung, Rahmenbedingungen, Befunde*. Wiesbaden: VS, Verl. für Sozialwiss, S. 23–43.
- Häcker, T. H. (2007). *Portfolio: ein Entwicklungsinstrument für selbstbestimmtes Lernen: Eine explorative Studie zur Arbeit mit Portfolios in der Sekundarstufe I*. Baltmannsweiler: Schneider Verl. Hohengehren.
- Häder, M., & Häder, S. (2019). Stichprobenziehung in der quantitativen Sozialforschung. In N. Baur, & J. Blasius (Hrsg.), *Handbuch Methoden der empirischen Sozialforschung*. Wiesbaden: Springer Fachmedien Wiesbaden, S. 333–348.
- Hafner, R. (2013). *Statistik: Für Sozial- und Wirtschaftswissenschaftler*. Wien: Springer.
- Hamann, P., & Erichson, B. (2000). *Marktforschung: 54 Tabellen*. Stuttgart: Lucius & Lucius.
- Hänssig, A. (2010). Portfolio-Arbeit in den Schulpraktischen Studien: Ein Praxisbeispiel. In , *Reflexion und Intervention : zur Theorie und Praxis schulpraktischer Studien*. Baltmannsweiler: Schneider-Verl. Hohengehren, S. 141–180.
- Hartig, J., & Klieme, E. (2006). Kompetenz und Kompetenzdiagnostik. In K. Schweizer (Hrsg.), *Leistung und Leistungsdiagnostik*. Berlin: Springer, S. 127–143.
- Hartig, J., & Klieme, E. (2008). Kompetenzkonzepte in den Sozialwissenschaften und im erziehungswissenschaftlichen Diskurs. In M. Prenzel, I. Gogolin, & H.-H. Krüger (Hrsg.), *Kompetenzdiagnostik: Zeitschrift für Erziehungswissenschaft*. Wiesbaden: VS Verlag für Sozialwissenschaften, S. 11–29.
- Hatton, N., & Smith, D. (1995). Reflection in teacher education: Towards definition and implementation. *Teaching and Teacher Education* 11, S. 33–49.
- Heinzel, F., & Brencher, G. (2008). Reflexionen von Studierenden zum Thema "Veränderte Kindheit" im Rahmen von Lerntagebüchern. In A. Hartinger (Hrsg.), *Veränderte Kindheit: Konsequenzen für die Lehrerbildung: Frau Professorin Dr. Maria Fölling-Albers zum 60.*

- Geburtstag*] ; [Beiträge einer Tagung, die im Oktober 2006 in Regensburg stattfand. Bad Heilbrunn: Klinkhardt, S. 33–49.
- Henderson, K., Napan, K., & Monteiro, S. (2004). Encouraging reflective learning: An online challenge. In R. Atkinson, C. McBeath, D. Jonas-Dwyer, & R. Phillips (Hrsg.), *Beyond the comfort zone: Proceedings of the 21st ASCILITE Conference*, S. 357–364.
- Herzog, W. (1995). *Reflexive Praktika in der Lehrerinnen- und Lehrerbildung*. Langnau, Emmental: Schweizerische Gesellschaft für Lehrerinnen- und Lehrerbildung (SGL).
- Hilscher, C. (2017). Kritisches Denken. <https://psylex.de/psychologie-lexikon/intelligenz/kritisches-denken.html>.
- Hilzensauer, W. (2008). *Theoretische Zugänge und Methoden zur Reflexion des Lernens. Ein Diskussionsbeitrag: Bildungsforschung, Bd. 2 (2008): Reflexives Lernen*.
- Hilzensauer, W. (2017). *Wie kommt die Reflexion in den Lehrberuf?* Dissertation. Universität Hamburg.
- Hubert, F., Meyer, F., & Lenzen, J. M. (2014). *Grundlagen der Varianzanalyse: Konzeption - Durchführung - Auswertung*. Wiesbaden: Springer Gabler.
- Hung, C.-Y. (2014). The Use of Weblogs in Citizenship Education: A Theoretical Analysis of the Integration of Weblogs and the English Citizenship Curriculum. *E-Learning and Digital Media* 11, 350-362. <http://dx.doi.org/10.2304/elea.2014.11.4.350>. Gesehen 5. Dezember 2019.
- Hunter, D. A. (2009). *A practical guide to critical thinking: Deciding what to do and believe*. Hoboken, N.J: John Wiley & Sons Inc.
- Hussy, W., Schreier, M., & Echterhoff, G. (2010). *Forschungsmethoden in Psychologie und Sozialwissenschaften für Bachelor*. Berlin, Heidelberg: Springer.
- Hüther, J. (2005). Neue Medien. In Hüther, J., Schorb (Hrsg.), *Grundbegriffe Medienpädagogik*. München: kopaed, S. 345–351.
- Jahn, D. (2013). Was es heißt, kritisches Denken zu fördern Ein pragmatischer Beitrag zur Theorie und Didaktik kritischen Nachdenkens. http://www.mediamanual.at/mediamanual/mm2/themen/kompetenz/mmt_1328_kritischesdenken_OK.pdf. Gesehen 5. Dezember 2019.
- Jahncke, H., Porath, J., Rebmann, K., Riebenbauer, E., & Stock, M. (2018). Vergleichende Analyse zweier Portfoliokonzepte zur Beförderung der (Selbst-)Reflexionsfähigkeit bei Studierenden der Wirtschaftspädagogik. *die hochschullehre* 4.

- Janczyk, M., & Pfister, R. (2013). *Inferenzstatistik verstehen: Von A wie Signifikanztest bis Z wie Konfidenzintervall*. Berlin, Heidelberg: Springer Berlin Heidelberg.
- Jasper, M. (2013). *Beginning reflective practice*. Australia: Cengage Learning.
- Jenert, T. (2008). Ganzheitliche Reflexion auf dem Weg zu Selbstorganisiertem Lernen. *Bildungsforschung* 5, S. 1–18.
- Johns, C. (1998). Opening the doors of perception. In C. Johns, & D. Freshwater (Hrsg.), *Transforming Nursing Through Reflective Practice*. New York, NY: John Wiley & Sons, S. 1–20.
- Karadag, M., Saritaş, S., & Erginer, E. (2009). Using the 'six thinking hats' model of learning in a surgical nursing class: Sharing the experience and student opinions. *Australian Journal of Advanced Nursing* 26, S. 59–69.
- Karmen, T., & Marju, M. (2015). Social Work Students' Thoughts on Self-Reflection: A Qualitative Study Based on Reflective Journaling. *International Journal of Humanities and Social Science* 5, S. 89–96.
- Kelle, U., & Kluge, S. (2010). *Vom Einzelfall zum Typus*. Wiesbaden: VS Verlag für Sozialwissenschaften.
- Keller, D. (2015). Welcher Post-Hoc-Test ist der Richtige? Statistik und Beratung. <https://statistik-und-beratung.de/2015/01/welcher-post-hoc-test-ist-der-richtige/>.
- Kember, D., Leung, D., Jones, A., & Loke, A. Y. (2000). Development of a questionnaire to measure the level of reflective thinking.: . *Assessment and Evaluation in Higher Education* 25, S. 380–395.
- Killus, D. (2009). Förderung selbstgesteuerten Lernens im Kontext lehrer- und organisationsbezogener Merkmale: Paralleltitel: The promotion of self-controlled learning within the context of teacher- and organization-related characteristics. *Zeitschrift für Pädagogik* 55, S. 130–150. <https://nbn-resolving.org/urn:nbn:de:0111-opus-42435>. Gesehen 5. Dezember 2019.
- King, P. M., & Kitchener, K. S. (1994). *Developing reflective judgment: Understanding and promoting intellectual growth and critical thinking in adolescents and adults*. San Francisco: Jossey-Bass.
- Kinsella, E. A. (2007). Technical rationality in Schön's reflective practice: Dichotomous or non-dualistic epistemological position. *Nursing Philosophy* 8, S. 102–113.
- Kittel, D., & Rollett, W. (2017). Entwickelt sich die professionelle Reflexionskompetenz von Lehrkräften durch die Teilnahme an einem berufsbegleitenden Weiterbildungsstudiengang? *Zeitschrift Hochschule und Weiterbildung*, 45-51.

- Kivunja, C. (2015a). Exploring the Pedagogical Meaning and Implications of the 4Cs " Super Skills " for the 21 st Century through Bruner's 5E Lenses of Knowledge Construction to Improve Pedagogies of the New Learning Paradigm. *Creative Education*, S. 224–239.
- Kivunja, C. (2015b). Using De Bono's Six Thinking Hats Model to Teach Critical Thinking and Problem Solving Skills Essential for Success in the 21st Century Economy. *Creative Education*, S. 380–391.
- Klieme, E. (2004). Was sind Kompetenzen und wie lassen sie sich messen? *Pädagogik* 56, S. 10–13.
- Klieme, E. (2007). *Zur Entwicklung nationaler Bildungsstandards: Eine Expertise*. Bonn, Berlin: Bundesministerium für Bildung und Forschung (BMBF).
- Kluge, K.-J., Weinreich, J., & Kluge, E. (o. J.). "Rote Heringe" im Unterricht (Hoch-) Begabter oder "Wenn wir testen und experimentieren, dann lernen wir und schreiten unaufhaltsam voran! Kritisches Denken ist der Motor der menschlichen Entwicklung und zum Weltfrieden". <http://www.criticalthinking-germany.de/Rote%20Heringe%20original.pdf>.
- KMK (2005). Einheitliche Prüfungsanforderungen in der Abiturprüfung - Geographie. Ständige Konferenz der Kultusminister der Länder in der Bundesrepublik Deutschland. https://www.kmk.org/fileadmin/veroeffentlichungen_beschluesse/1989/1989_12_01-EPA-Geographie.pdf.
- KMK und BMZ (2015). *Orientierungsrahmen für den Lernbereich globale Entwicklung im Rahmen einer Bildung für nachhaltige Entwicklung: Ein Beitrag zum Weltaktionsprogramm „Bildung für nachhaltige Entwicklung“*.
- Kolb, D. A. (1984). *Experiential learning: Experience as the source of learning and development*. Upper Saddle River, NJ: Prentice-Hall.
- Kölle, B. (2017). *Rahmenlehrplan Teil C - Geographie: Themen und Inhalte des Berliner Unterrichts im Überblick*. Berlin: Senatsverwaltung für Bildung Jugend und Familie.
- Koller, P. J., & Alpar, P. (2008). Die Bedeutung privater Weblogs für das Issue-Management in Unternehmen. In P. Alpar, & S. Blaschke (Hrsg.), *Web 2.0 - eine empirische Bestandsaufnahme*. Wiesbaden: Vieweg+Teubner / GWV Fachverlage GmbH Wiesbaden, S. 17–52.
- König, C., Stahl, M., & Wiegand, E. (2014). *Soziale Medien: Gegenstand und Instrument der Forschung*. Wiesbaden: Springer VS.
- Kopp, D. von (2015). *Focusing: Die Sprache der Intuition*. Wiesbaden Germany: Springer.
- Korthagen, F. A. J. (2002). Eine Reflexion über Reflexion. In F. A. J. Korthagen (Hrsg.), *Schulwirklichkeit und Lehrerbildung: Reflexion der Lehrertätigkeit*, S. 55–73.

- Kraak, B. (o. J.). Kritisches Denken als Bildungsaufgabe. Institut für Leistungsentwicklung.
<https://www.hochbegabtenhilfe.de/kritisches-denken-als-bildungsaufgabe/>.
- Krapp, A. (2010). Die Bedeutung von Interessen für die Lernmotivation und das schulische Lernen – eine Einführung. In I. Hemmer (Hrsg.), *Schülerinteresse an Themen, Regionen und Arbeitsweisen des Geographieunterrichts: Ergebnisse der empirischen Forschung und deren Konsequenzen für die Unterrichtspraxis*. Weingarten: HGD, S. 9–26.
- Krippendorff, K. (1980). *Content analysis: An introduction to its methodology*. Beverly Hills: Sage Publications.
- Kroeber-Riel, W., & Weinberg, P. (1999). *Konsumentenverhalten*. München: Vahlen.
- Krosnick, J. A., & Presser, S. (2010). Question and Questionnaire Design. In P. V. Marsden, & James D. (Hrsg.), *Handbook of survey research*. Bingley: Emerald, S. 264–313.
- Kuckartz, U. (2012). *Qualitative Inhaltsanalyse: Methoden, Praxis, Computerunterstützung*. Weinheim: Beltz-Juventa.
- Kuckartz, U., Rädiker, S., Ebert, T., & Schehl, J. (2013). *Statistik: Eine verständliche Einführung*. Wiesbaden: Springer VS.
- Künzli, C., Bertschy, F., & Buchs, C. (2013). Bildung für eine Nachhaltige Entwicklung – Rolle und Aufgaben von Lehrpersonen. In BNE-Konsortium COHEP (Hrsg.), *Didaktische Grundlagen zur Bildung für Nachhaltige Entwicklung in der Lehrerinnen- und Lehrerbildung*. Zürich/Fribourg, S. 60–64.
- LaBoskey, V. K. (1993). A conceptual framework for reflection in preservice teacher education. In J. Calderhead, & P. Gates (Hrsg.), *Conceptualizing Reflection in Teacher Development*. London: Washington Falmer Press, S. 23–38.
- Lamberti, J. (2001). *Einstieg in die Methoden empirischer Forschung: Planung, Durchführung und Auswertung empirischer Untersuchungen*. Tübingen: Dgvt-Verl.
- Lamnek, S., & Krell, C. (2016). *Qualitative Sozialforschung*. Weinheim: Beltz.
- Landmann, M., & Schmitz, B. (2007). Die Kombination von Trainings mit standardisierten Tagebüchern: Angeleitete Selbstbeobachtung als Möglichkeit der Unterstützung von Trainingsmaßnahmen. In M. Landmann, & B. Schmitz (Hrsg.), *Selbstregulation erfolgreich fördern. Praxisnahe Trainingsprogramme für effektives Lernen*. Stuttgart: Kohlhammer, S. 151–163.
- Leidinger, M. (2014). *Förderung von Strategien selbstregulierten Lernens und deren Einfluss auf die schulische Leistung sowie die Selbstwirksamkeitsüberzeugungen von Schülern im Primarbereich. Implementation einer Lernumgebung in den regulären Unterricht der vierten Klassenstufe*. Saarbrücken: Saarländische Universitäts- und Landesbibliothek.

- Leonhard, T. (2008). *Professionalisierung in der Lehrerbildung: Eine explorative Studie zur Entwicklung professioneller Kompetenzen in der Lehrererstausbildung*. Zugl.: Heidelberg, Pädag. Hochsch., Diss. Berlin: Logos.
- Levine, R., & Vogel, R. (2000). *Das Cluetrain-Manifest: 95 Thesen für die neue Unternehmenskultur im digitalen Zeitalter*. München: Econ-Verl.
- Lisch, R., & Kriz, J. (1978). *Grundlagen und Modelle der Inhaltsanalyse: Bestandsaufnahme und Kritik*. Reinbek bei Hamburg: Rowohlt.
- Mann, K., Gordon, J., & Macleod, A. (2007). Reflection and Reflective Practice in Health Professions Education: A Systematic Review. *Advances in health sciences education : theory and practice*, S. 595–621.
- Mansvelder-Longayroux, D. D., Beijaard, D., & Verloop, N. (2007). The Portfolio as a Tool for Stimulating Reflection by Student Teachers. *Teaching & Teacher Education: An International Journal of Research and Studies* 23, 47-62.
<http://dx.doi.org/10.1016/j.tate.2006.04.033>. Gesehen 5. Dezember 2019.
- Marrapodi, J. (2003). *Critical thinking and creativity: An overview and comparison of the theories*. Providence.
- Matzen, N. (2014). *Onlinejournalismus*. Berlin: UVK Verlagsgesellschaft.
- Maxwell, J. A., & Loomis, D. M. (2003). Mixed Methods Design: An alternative approach. In A. Tashakkori, & C. Teddlie (Hrsg.), *Handbook of mixed methods in social & behavioral research*. Thousand Oaks, Calif.: Sage, S. 241–272.
- Mayring, P. (2010). *Qualitative Inhaltsanalyse: Grundlagen und Techniken*. Weinheim, Basel: Beltz.
- Mayring, P. (2015). *Qualitative Inhaltsanalyse: Grundlagen und Techniken*. Weinheim: Beltz.
- McMillan, J. H., & Schumacher, S. (2006). *Research in education: Evidence-based inquiry*. Boston: Pearson/Allyn and Bacon.
- Menold, N., & Bogner, K. (2016). *Design of Rating Scales in Questionnaires (Version 2.0)*. Mannheim.
- Mertens, D. M. (2005). *Research and evaluation in education and psychology: Integrating diversity with quantitative, qualitative, and mixed methods*. Thousand Oaks, Calif.: Sage Publications.
- Metz, M. (2016). *Nachhaltiges Konsumhandeln - Relevanz von Reflexionsfähigkeit im Rahmen von Bildung für nachhaltige Entwicklung*. Dissertation. VVB LAUFERSWEILER Verlag.

- Meyer, H. (1987). *Unterrichtsmethoden*. Frankfurt am Main: Cornelsen Scriptor.
- Mezirow, J. (1981). A critical theory of adult learning and education. *Adult Education Quarterly* 32, S. 3–24.
- Michelsen, G. (2008). Nachhaltigkeit als Herausforderung. Mentalitätswandel ist die Aufgabe von Bildung. *Erziehung und Wissenschaft*, S. 6–9.
- Miles, M. B., & Huberman, A. M. (2009). *Qualitative data analysis: An expanded sourcebook*. Thousand Oaks, Calif.: Sage.
- Mohd Razali, N., & Yap, B. (2011). Power Comparisons of Shapiro-Wilk, Kolmogorov-Smirnov, Lilliefors and Anderson-Darling Tests. *J. Stat. Model. Analytics* 2.
- Moon, J. A. (2004). *A handbook of reflective and experiential learning: Theory and practice*. London, New York: RoutledgeFalmer.
- Moon, J. A. (2006). *Learning journals: A handbook for reflective practice and professional development*. London: Routledge.
- Moon, J. A. (2007). *Critical Thinking: An Exploration of Theory and Practice*. Routledge.
- Moon, J. A. (2013). *Reflection in Learning and Professional Development: Theory and Practice*. Hoboken: Taylor and Francis.
- Morgan, D. L. (1998). Practical strategies for combining qualitative and quantitative methods: applications to health research. *Qualitative health research*, S. 362–376.
- Morse, J. (1991). Approaches to Qualitative-Quantitative Methodological Triangulation. *Nursing Research* 40, S. 120–123.
- Mühlhausen, U. (Hrsg.). (2006). *Unterrichten lernen mit Gespür: Szenarien für eine multimedial gestützte Analyse und Reflexion von Unterricht*. Baltmannsweiler: Schneider Verl. Hohengehren.
- Mulder, M. (2014). Conceptions of Professional Competence. In S. Billett, C. Harteis, & H. Gruber (Hrsg.), *International Handbook of Research in Professional and Practice-based Learning*. Dordrecht: Springer, S. 107–137.
- Nestvogel, R. (2002). Zum Verhältnis von „Interkulturellem Lernen“, „Globalen Lernen“ und „Bildung für nachhaltige Entwicklung“. In C. Wulf, & C. Merkel (Hrsg.), *Globalisierung als Herausforderung der Erziehung: Theorien, Grundlagen, Fallstudien*. Münster: Waxmann, S. 31–44.
- O'Reilly, T. (2005). What Is Web 2.0: Design Patterns and Business Models for the Next Generation of Software. <http://www.oreilly.com/pub/a/web2/archive/what-is-web-20.html>. Gesehen 5. Dezember 2019.

- Ohl, U., & Klebel, C. (2012). Simulationsspiele in der Bildung für nachhaltige Entwicklung (BNE). In M. Pingold, & R. Uphues (Hrsg.), *Jenseits des Nürnberger Trichters Ideen für einen zukunftsorientierten; Geographieunterricht : Tagungsband zum 15. Bayerischen Schulgeographentag, 12. - 14.10.2012, Nürnberg*. Erlangen: Friedrich-Alexander-Universität, 57-61 und 143-144.
- Oner, D., & Adadan, E. (2011). Use of Web-Based Portfolios as Tools for Reflection in Pre-service Teacher Education. *Journal of Teacher Education*, S. 477–492.
- Overwien, B. (2009). Globalisierung und Globales Lernen. In T. Lucker, & O. Kölsch (Hrsg.), *Naturschutz und Bildung für nachhaltige Entwicklung - Fokus Globales Lernen: Ergebnisse des F+E-Vorhabens "Bildung für nachhaltige Entwicklung (BNE) - Positionierung des Naturschutzes"*. Bonn: Bundesamt für Naturschutz, S. 63–80.
- Parducci, A. (1983). Category ratings and the relational character of judgment. In H.-G. Geisler (Hrsg.), *Modern issues in perception*. Berlin: Dt. Verl. d. Wiss, S. 262–282.
- Pross, H. (1970). *Publizistik: Thesen zu einem Grundcolloquium*. Neuwied: Luchterhand.
- Quinn, F. M. (2000). *The principles and practice of nurse education*. Cheltenham, U.K.: Stanley Thornes.
- Rahm, S., & Lunkenbein, M. (2014). Anbahnung von Reflexivität im Praktikum: Empirische Befunde zur Wirkung von Beobachtungsaufgaben im Grundschulpraktikum. In K.-H. Arnold, A. Gröschner, & T. Hascher (Hrsg.), *Schulpraktika in der Lehrerbildung: Theoretische Grundlagen, Konzeptionen, Prozesse und Effekte = Pedagogical field experiences in teacher education*. Münster, New York, NY: Waxmann, S. 237–256.
- Rana, D., & Upton, D. (2013). *Psychology for Nurses*. Hoboken: Taylor and Francis.
- Rasch, B., Friese, M., Hofmann, W., & Naumann, E. (2014). *Quantitative Methoden 1: Einführung in die Statistik für Psychologen und Sozialwissenschaftler*. Berlin: Springer.
- Reinders, H., Ditton, H., Gräsel, C., & Gniewosz, B. (2015). *Empirische Bildungsforschung: Strukturen und Methoden*. Wiesbaden: Springer Fachmedien Wiesbaden.
- Reinmann, G. (2005). *Erfahrungswissen erzählbar machen: Narrative Ansätze für Wirtschaft und Schule*. Lengerich u.a: Pabst Science Publishers.
- Richardson, W. (2011). *Wikis, Blogs und Podcasts: Neue und nützliche Werkzeuge für den Unterricht*. Überlingen: TibiaPress.
- Rieckmann, M. (2010). *Die globale Perspektive der Bildung für eine nachhaltige Entwicklung*. Zugl.: Lüneburg, Univ., Diss., 2010. BWV, Berliner Wiss.-Verl. Berlin.

- Rieckmann, M. (2013). Schlüsselkompetenzen für eine nachhaltige Entwicklung. *POLIS*, S. 11–13.
- Rieckmann, M. (2016). Kompetenzentwicklungsprozesse in der Bildung für nachhaltige Entwicklung erfassen: Überblick über ein heterogenes Forschungsfeld. In M. Rieckmann, & M. Barth (Hrsg.), *Empirische Forschung zur Bildung für nachhaltige Entwicklung - Themen, Methoden und Trends*. Opladen: Verlag Barbara Budrich, S. 89–109.
- Rinschede, G. (1999). Medien. In D. Böhn (Hrsg.), *Didaktik der Geographie - Begriffe*. München: Oldenbourg, S. 101–102.
- Röbken, H., & Wetzel, K. (2016). *Qualitative und quantitative Forschungsmethoden*. Oldenburg.
- Rodgers, C. (2002). Defining Reflection: Another Look at John Dewey and Reflective Thinking. *Teachers College Record* 4, S. 842–866. https://www.researchgate.net/publication/240645823_Defining_Reflection_Another_Look_at_John_Dewey_and_Reflective_Thinking. Gesehen 5. Dezember 2019.
- Rudolf, M., & Müller, J. (2012). *Multivariate Verfahren: Eine praxisorientierte Einführung mit Anwendungsbeispielen in SPSS*. Göttingen, Bern, Wien, Paris: Hogrefe.
- Rychen, D. S. (2003). Key competencies: Meeting important challenges in life. In D. S. Rychen, & L. H. Salganik (Hrsg.), *Key competencies for a successful life and a well-functioning society*. Cambridge, Mass.: Hogrefe & Huber, S. 63–108.
- Sachs, L. (1997). *Angewandte Statistik: Anwendung statistischer Methoden*. Berlin: Springer.
- Saldern, M. v. (1991). *Erziehungswissenschaft und neue Systemtheorie*. Berlin: Duncker & Humblot.
- Sandelowski, M., Voils, C. I., & Knafl, G. (2009). On Quantitizing. *Journal of mixed methods research*, S. 208–222.
- Sauter, W., & Sauter, S. (2013). *Workplace Learning: Integrierte Kompetenzentwicklung mit kooperativen und kollaborativen Lernsystemen*. Berlin: Springer Gabler.
- Schärli, M. (2013). Ergebnisse der Metakognition im E-Portfolio. *Pflegewissenschaft*, S. 466–480.
- Schellmann, B., Baumann, A., Gläser, M., Kegel, T., & Verlag Europa-Lehrmittel Nourney, Vollmer GmbH & Co (2017). *Handbuch Medien: Medien verstehen - gestalten - produzieren*. Verlag Europa-Lehrmittel Nourney, Vollmer GmbH & Company KG.

- Scheunpflug, A. (2001). Der Beitrag der Entwicklungspädagogik zum Agenda-Prozeß. In O. Herz, H. Seybold, & G. Strobl (Hrsg.), *Bildung für nachhaltige Entwicklung: Globale Perspektiven und neue Kommunikationsmedien*. Wiesbaden: VS Verlag für Sozialwissenschaften, S. 87–100.
- Scheunpflug, A., & Asbrand, B. (2006). Global education and education for sustainability. *Environmental Education Research*, S. 33–46.
- Schmal, J. (2017). Grundsätze der Pädagogik in Gesundheitsfachberufen. In J. Schmal (Hrsg.), *Unterrichten und Präsentieren in Gesundheitsfachberufen*. Berlin, Heidelberg: Springer Berlin Heidelberg, S. 17–45.
- Schmal, J. (2018). Reflexionsfähigkeit in Pflegeausbildung und -studium fördern. *Pflegezeitschrift*, S. 52–55.
- Schneider, A. (2013). *Geographiedidaktische Reflexivität: Ostdeutsche Mobilitätsfragen im zweiten Blick*. Zugl.: Jena, Univ., Diss., 2012. Berlin: Lit.
- Schockemöhle, J., & Schrüfer, G. (2015). Bildung für nachhaltige Entwicklung - Leitbild für das Lernen in verschiedenen Ernährungssituationen. In J. Schockemöhle, & M. Stein (Hrsg.), *Nachhaltige Ernährung lernen in verschiedenen Ernährungssituationen: Handlungsmöglichkeiten in pädagogischen und sozialpädagogischen Einrichtungen*. Bad Heilbrunn: Klinkhardt, S. 46–59.
- Schön, D. A. (1983). *The reflective practitioner: How professionals think in action*. New York: Basic Books.
- Schwarzer, G. (2001). Forschungsanträge verfassen – Ein praktischer Ratgeber für Sozialwissenschaften. *Zeitschrift für Internationale Beziehungen* 8, S. 141–156.
- Smyth, W. J. (1992). Teachers' work and the politics of reflection. *American Educational Research Journal*, S. 267–300.
- Stangl, W. (2019). Stichwort: 'erfahrungsbasiertes Lernen'. Online Lexikon für Psychologie und Pädagogik. <https://lexikon.stangl.eu/14476/erfahrungsbasiertes-lernen/>. Gesehen 5. Dezember 2019.
- Steinskog, D., & Kvamstø, N. (2007). A Cautionary Note on the Use of the Kolmogorov–Smirnov Test for Normality. *Monthly Weather Review - MON WEATHER REV*, S. 1151–1157.
- Stocker, C. (2007). Zwischen Wunsch und Wirklichkeit: Weblogs im Hochschulunterricht. In U. Dittler, M. Kindt, & C. Schwarz (Hrsg.), *Online-Communities als soziale Systeme: Wikis, Weblogs und Social Software im E-Learning*. Münster, New York, München, Berlin: Waxmann, S. 97–115.

- Tann, S. (1993). Eliciting student teachers' personal theories. In J. Calderhead, & P. Gates (Hrsg.), *Conceptualizing Reflection in Teacher Development*. London: Falmer Press, S. 53–69.
- Tashakkori, A., & Teddlie, C. (1998). *Mixed methodology: Combining qualitative and quantitative approaches*. Thousand Oaks, Calif.: Sage Publications.
- Tashakkori, A., & Teddlie, C. (2003). The Past and Future of Mixed Methods Research: From Data Triangulation to Mixed Model Designs. In A. Tashakkori, & C. Teddlie (Hrsg.), *Handbook of mixed methods in social & behavioral research*. Thousand Oaks, Calif.: Sage, S. 671–701.
- Tawanwongsri, W., & Phenwan, T. (2019). Reflective and feedback performances on Thai medical students' patient history- taking skills. *BMC Medical Education*.
- Tulodziecki, G., Herzig, B., & Blömeke, S. (2004). *Gestaltung von Unterricht: Eine Einführung in die Didaktik*. Bad Heilbrunn: Klinkhardt.
- Tzeng, W.-S., Kuo, K.-M., Talley, P. C., Chen, H.-C., & Wang, J.-J. (2015). Do ePortfolios Contribute to Learners' Reflective Thinking Activities? : A Preliminary Study of Nursing Staff Users. *Journal of medical systems* 39, S. 1–10.
- Ulich, D., Hausser, K., Mayring, P., Strehmel, P., Kandler, M., & Degenhardt, B. (1985). *Psychologie der Krisenbewältigung: Eine Längsschnittuntersuchung mit arbeitslosen Lehrern*. Weinheim: Beltz.
- van den Boom, G., Paas, F., & van Merriënboer, J. J. G. (2007). Effects of Elicited Reflections Combined with Tutor or Peer Feedback on Self-Regulated Learning and Learning Outcomes. *Learning and Instruction* 17, 532-548.
- van den Boom, G., Paas, F., van Merriënboer, Jeroen J. G., & van Gog, T. (2004). Reflection prompts and tutor feedback in a webbased learning environment: Effects on students' self-regulated learning competence. *Computers in Human Behavior*.
- van Manen, M. (1977). Linking ways of knowing with ways of being practical. *Curriculum Inquiry* 6, S. 205–228.
- Vare, P., & Scott, W. (2007). Learning for a Change. *Journal of Education for Sustainable Development*, S. 191–198.
- Vock, M., & Gronostaj, A. (2017). *Umgang mit Heterogenität in Schule und Unterricht*. Berlin: Friedrich-Ebert-Stiftung Abt. Studienförderung.
- Vogl, S. (2017). Quantifizierung: Datentransformation von qualitativen Daten in quantitative Daten in Mixed-Methods-Studien. In , *Mixed methods*. Wiesbaden: Springer VS, S. 287–312.

- Wals, A. (2011). Learning Our Way to Sustainability. *Journal of Education for Sustainable Development*, S. 177–186.
- Weidemann, B. (2002). Multicodierung und Multimodalität im Lernprozess. In L. J. Issing, & P. Klimsa (Hrsg.), *Information und Lernen mit Multimedia und Internet: Lehrbuch für Studium und Praxis*. Weinheim: Beltz PVU, S. 45–64.
- Weinert, F. E. (2002). Vergleichende Leistungsmessung in Schulen - eine umstrittene Selbstverständlichkeit. In F. E. Weinert (Hrsg.), *Leistungsmessungen in Schulen*. Weinheim: Beltz, S. 17–31.
- Wild, E., & Möller, J. (2009). *Pädagogische Psychologie*. Berlin, Heidelberg: Springer-Verlag Berlin Heidelberg.
- Winter, S. (2000). Quantitative vs. Qualitative Methoden. http://nosnos.synology.me/MethodenlisteUniKarlsruhe/imihome.imi.uni-karlsruhe.de/nquantitative_vs_qualitative_methoden_b.html. Gesehen 5. Dezember 2019.
- Wirtz, M. A. (2014). *Dorsch Lexikon der Psychologie*. Bern: Huber.
- Wittkowski, J. (2013). *Das Interview in der Psychologie: Interviewtechnik und Codierung von Interviewmaterial*. Berlin: Springer.
- Wolf, W. (1995). Qualitative versus quantitative Forschung. In E. König, & P. Zedler (Hrsg.), *Bilanz qualitativer Forschung. 1. Grundlagen qualitativer Forschung*. Weinheim: Deutscher Studien Verl., S. 309–329.
- Wyss, C. (2008). Zur Reflexionsfähigkeit und -praxis der Lehrperson. *Bildungsforschung* 5.
- Wyss, C. (2013). *Unterricht und Reflexion: Eine mehrperspektivische Untersuchung der Unterrichts- und Reflexionskompetenz von Lehrkräften*. Münster: Waxmann.
- Zander, S., & Brünken, R. (2006). Lernen mit neuen Medien in den Naturwissenschaften. In A. Pitton (Hrsg.), *Lehren und Lernen mit neuen Medien: Gesellschaft für Didaktik der Chemie und Physik, Band 26 ; (32.) Jahrestagung der GDCh in Paderborn 2005*. Berlin: LIT-Verl., S. 15–17.
- Zdarzil, H. (1978). *Pädagogische Anthropologie*. Zugl.: Bonn, Univ., Habil.-Schr., 1970. Graz: Verl. Styria.
- Zerfaß, A., & Boelter, D. (2005). *Die neuen Meinungsmacher: Weblogs als Herausforderung für Kampagnen, Marketing, PR und Medien*. Graz: Nausner und Nausner.
- Zimmermann, M., & Welzel, M. (2008). Reflexionskompetenz – ein Schlüssel zur naturwissenschaftlichen Frühförderkompetenz. *Perspektiven zur pädagogischen Professionalisierung* 74, S. 29–36.

Abbildungsverzeichnis

Abbildung 1	Positionierung von Blogs innerhalb der Medienklassifikation (eigener Entwurf; in Anlehnung an Pross (1970), Faßler (1997), Brendel und Schröder (2013)).....	47
Abbildung 2	Überblick über übergeordnetes Forschungskonzept für diese Studie.....	65
Abbildung 3	Ablaufmodell der qualitativen Inhaltsanalyse in Form der skalierenden Strukturierung nach Mayring (2015).....	81
Abbildung 4	Kodierter Text mit Anwendung beider Kategoriensysteme zum Thema Einkaufen von P3.....	91
Abbildung 5	Scorewerte erster und letzter Beitrag	106
Abbildung 6	Reflexivitätsverteilung erster und letzter Beitrag	107
Abbildung 7	Reflexionsverteilung erster und letzter Beitrag	108
Abbildung 8	Scorewerte der Textblöcke 1, 2 und 3.....	109
Abbildung 9	Reflexivitätsverteilung der Textblöcke 1, 2 und 3	110
Abbildung 10	Reflexionsverteilung der Textblöcke 1, 2 und 3.....	112
Abbildung 11	Fachliche Nähe bei Definition vor und nach Intervention.....	114
Abbildung 12	Teilbereiche des Reflektierens in Definition vor und nach Intervention.....	116
Abbildung 13	Gesamtscore im Reflektieren – erster und letzter Beitrag	118
Abbildung 14	Reflexionsscore – erster und letzter Beitrag.....	119
Abbildung 15	Reflexivitätsscore – erster und letzter Beitrag	120
Abbildung 16	Maximal erreichte Reflexionsstufe – erster und letzter Beitrag.....	121
Abbildung 17	Maximal erreichte Reflexivitätsstufe – erster und letzter Beitrag	122
Abbildung 18	Am häufigsten erreichte Reflexionsstufe – erster und letzter Beitrag	123
Abbildung 19	Am häufigsten erreichte Reflexivitätsstufe – erster und letzter Beitrag	124
Abbildung 20	Gesamtscore im Reflektieren - vor, zwischen und nach Teilinterventionen	125
Abbildung 21	Reflexionsscore - vor, zwischen und nach Teilinterventionen	127
Abbildung 22	Reflexivitätsscore - vor, zwischen und nach Teilinterventionen.....	129
Abbildung 23	Maximal erreichte Reflexionsstufe - vor, zwischen und nach Teilinterventionen	130
Abbildung 24	Maximal erreichte Reflexivitätsstufe - vor, zwischen und nach Teilinterventionen	132
Abbildung 25	Am häufigsten erreichte Reflexionsstufe - vor, zwischen und nach Teilinterventionen	133
Abbildung 26	Am häufigsten erreichte Reflexivitätsstufe - vor, zwischen und nach Teilinterventionen	135

Tabellenverzeichnis

Tabelle 1	Kernkompetenzen des Bereichs Bewerten des Orientierungsrahmens für den Lernbereich Globale Entwicklung im Rahmen einer BNE (KMK und BMZ 2015, S. 96), eigene Überarbeitung.....	21
Tabelle 2	Niveaumodelle hinsichtlich Reflektierfähigkeiten.....	24
Tabelle 3	Strukturmodell der Reflektierfähigkeit	25
Tabelle 4	Übersicht der Reflexionsmodelle, die als Grundlage für Brendels Stufenmodell (2017) zur Bestimmung von Reflexionsfähigkeit dienen	26
Tabelle 5	Stufenmodell reflexiven Denkens im Kontext Globalen Lernens nach Brendel (2017, S. 251).....	28
Tabelle 6	Stufenmodell nach Zimmermann und Welzel (2008, S. 34) zur Bestimmung von Reflexivitätsfähigkeit.....	30
Tabelle 7	Reflektierarten nach Hatton und Smith (1995, S. 48 f.)	31
Tabelle 8	Phasen des <i>reflective cycle</i> nach Gibbs (2013, S. 49 f.)	38
Tabelle 9	Übersicht der behandelten Themen im Blog	77
Tabelle 10	Kategoriensystem und Kodierleitfaden für Reflexionsfähigkeiten	83
Tabelle 11	Kategoriensystem und Kodierleitfaden für Reflexivitätsfähigkeit	86
Tabelle 12	Grundlage Scoreberechnung für Reflexionsfähigkeit	95
Tabelle 13	Grundlage Scoreberechnung für Reflexivitätsfähigkeit.....	95
Tabelle 14	Übersicht P5	105
Tabelle 15	Übersicht Testpersonen.....	113
Tabelle 16	Übersicht Gruppeneinteilung in Hinblick auf fachliche Nähe bei Definition ..	115
Tabelle 17	Übersicht Gruppeneinteilung in Hinblick auf Teilbereiche des Reflektierens in Definition	116
Tabelle 18	Übersicht Gruppeneinteilung in Hinblick auf Entwicklung der Reflektierfähigkeit vom ersten zum letzten Beitrag.....	117
Tabelle 19	Übersicht Gruppeneinteilung in Hinblick auf die Entwicklung des Gesamtscores im Reflektieren	126
Tabelle 20	Übersicht Gruppeneinteilung in Hinblick auf die Entwicklung des Reflexionsscores	128
Tabelle 21	Übersicht Gruppeneinteilung in Hinblick auf die Entwicklung des Reflexivitätsscores.....	129
Tabelle 22	Übersicht Gruppeneinteilung in Hinblick auf die Entwicklung der maximal erreichten Reflexionsstufe	131
Tabelle 23	Übersicht Gruppeneinteilung in Hinblick auf die Entwicklung der maximal erreichten Reflexivitätsstufe.....	132

Tabelle 24	Übersicht Gruppeneinteilung in Hinblick auf Entwicklung der am häufigsten erreichten Reflexionsstufe	134
Tabelle 25	Übersicht Gruppeneinteilung in Hinblick auf die Entwicklung der am häufigsten erreichten Reflexivitätsstufe.....	135
Tabelle 26	Gruppeneinteilung hinsichtlich der Entwicklung der Reflexivitätsfähigkeit nach erster Teilintervention.....	137
Tabelle 27	Gruppeneinteilung hinsichtlich der Entwicklung der Reflexionsfähigkeit nach erster Teilintervention.....	137
Tabelle 28	Gruppeneinteilung hinsichtlich der Entwicklung der Reflexivitätsfähigkeit nach zweiter Teilintervention.....	138
Tabelle 29	Gruppeneinteilung hinsichtlich der Entwicklung der Reflexionsfähigkeit nach zweiter Teilintervention.....	138
Tabelle 30	Effektstärken auf die Variablen vom ersten zum letzten Beitrag aller drei Faktorstufen.....	143
Tabelle 31	Effektstärken durch Reflektiermethoden	147
Tabelle 32	Effektstärken nach d Cohen vom ersten zum letzten Beitrag beider Experimentalgruppen.....	150
Tabelle 33	Effektstärken nach d Cohen vom ersten zum letzten Textblock beider Experimentalgruppen.....	151
Tabelle 34	Effektstärken hinsichtlich des Geschlechts	153

Abkürzungsverzeichnis

ANOVA = Varianzanalyse

BMZ = Bundesministerium für wirtschaftliche Zusammenarbeit und Entwicklung

BNE = Bildung für nachhaltige Entwicklung

DGfG = Deutsche Gesellschaft für Geographie

KMK = Kultusministerkonferenz

Anhang

- a. Arbeitsblatt für SuS mit Arbeitsdefinition des Begriffs *Reflektieren*
- b. Fragebogenvorlage zu Beginn und am Ende des Blogprojekts
- c. Arbeitsblatt Gibbs Reflexivitätszyklus
- d. Arbeitsblatt Hütemethode von de Bono
- e. Exemplarischer Blogbeitrag von P18
- f. Gegenüberstellung Zimmermann und Welzels Stufenmodell (2008) und Kategoriensystem zur Bestimmung der Stufen der Reflexivitätsfähigkeit
- g. Einzelauswertungen im Rahmen der qualitativen Inhaltsanalyse (P1 - P4)
- h. Deskriptive Statistik

Weiteres Material erhältlich auf Nachfrage bei der Autorin (anja.hager@uni-bayreuth.de):

- Material der Blogmappe
- Material der Lehrermappe inkl. Themenpool
- Blogbeiträge der Schülerinnen und Schüler
- Qualitative Auswertung der Blogbeiträge
- Restliche, hier nicht abgedruckte Einzelauswertungen der Probanden

a. Arbeitsblatt für SuS mit Arbeitsdefinition des Begriffs *Reflektieren*Was bedeutet „Reflektieren“?

Reflektieren bedeutet **Sachverhalte, Wissen, eigene und fremde Verhaltensweisen und Einstellungen kritisch zu hinterfragen** und diese daraufhin zu überprüfen, warum sie ihre Gültigkeit und ihren Wert besitzen (oder auch nicht).

Unter Reflektieren kann also ein **tiefgreifendes Nachdenken über bestimmte Themen und Handlungen** verstanden werden.

In euren Blogbeiträgen sollt ihr immer sowohl über das Thema selbst als auch über euer Verhalten diesbezüglich reflektieren!



pixabay.com/de/homer-simpsons-cartoon-comic-155238/

b. Fragebogenvorlage zu Beginn und am Ende des Blogprojekts

Bitte fülle diese Umfrage aus. Es gibt kein richtig oder falsch bei den Antworten! Vielen Dank fürs Mitmachen!

*** Erforderlich**

1. Wie lautet dein Username für den Blog? *

2. Dein Geschlecht? *

Markieren Sie nur ein Oval.

☐ weiblich

☐ männlich

3. Was bedeutet "Reflektieren" für dich? Erkläre den Begriff in 1-3 Sätzen! *

4. Bitte wähle jeweils die Antwort aus, die am ehesten zutrifft: *

Markieren Sie nur ein Oval pro Zeile.

	Stimmt voll und ganz	Stimmt eher schon	Stimmt eher nicht	Stimmt absoolut nicht
Gibbs-Reflexionszyklus als Methode war verständlich.	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Die Hüte-Methode war verständlich.	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Gibbs Reflexionszyklus half mir beim Reflektieren über mein Verhalten zu einem bestimmten Thema.	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Die Hüte-Methode half mir beim Reflektieren über ein bestimmtes Thema.	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>

Frage 4 ist nur im Fragebogen am Ende des Blogprojekts Bestand-

5. Bitte wähle die Antwort aus, die am ehesten auf dich zutrifft. *

Markieren Sie nur ein Oval pro Zeile.

	Stimmt voll und ganz	Stimmt eher schon	Stimmt eher nicht	Stimmt absolut nicht
Ich weiß wie Bloggen funktioniert.	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Bloggen finde ich sinnvoll.	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Bloggen macht bestimmt Spaß.	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Ich schreibe gerne Texte.	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Ich arbeite gerne mit digitalen Medien.	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Ich treffe Entscheidungen eher spontan ohne viel Nachzudenken.	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Ich habe immer eine Erklärung für mein Verhalten.	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Ich denke oft über mein Verhalten nach, um die Gründe dafür verstehen zu können.	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Ich akzeptiere andere Ansichten oder Meinungen bei komplexen Themen.	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Manchmal frage ich mich wie andere etwas tun und versuche eine bessere Lösung zu finden.	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Ich denke oft in meiner Freizeit über Unterrichtsinhalte nach.	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Neue Unterrichtsinhalte helfen mir, meine Ansichten zu überdenken.	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Checklisten helfen mir bei der Bearbeitung von Aufgaben.	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Bei Diskussionen bringe ich mich gerne ein.	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Diskussionsrunden helfen mir, ein Thema besser zu verstehen.	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Mit Verallgemeinerungen komme ich gut zurecht.	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>

c. Arbeitsblatt Gibbs Reflexivitätszyklus

Gibbs Reflexivitätszyklus zum Reflektieren über eines deiner Verhaltensgewohnheiten bezüglich eines bestimmten Themas

Diese Schritte helfen dir, eine gute Reflexion über dein Verhalten zu schreiben:



Die folgenden Fragen helfen dir, beim Schreiben einer Reflexion über eine deiner Verhaltensgewohnheiten zu einem bestimmten Thema. Die Fragen sind Bestandteile der einzelnen Phasen von Gibbs Reflexivitätszyklus.

Hilfsfragen für die einzelnen Phasen:**1. Beschreibung**

Beschreibe eine deiner gewöhnlichen Verhaltensweisen / Handlungen (in Bezug auf das entsprechende Thema) detailliert! Fragen, die dir dabei helfen:

Wann und wo passiert es?
Warum passiert es?
Wer ist daran beteiligt?
Was geschieht?
Was tust du?
Was machen andere Leute dabei?
Was ist das Ergebnis dieser Verhaltensweise / Handlung?

2. Emotionen

Wie fühlst du dich während der Situation?
Was beeinflusst deine Gefühle dabei (andere Menschen, Vorwissen...)?
Inwieweit beeinflussen deine Gefühle dein Verhalten / deine Handlung?
Was denkst du, wie sich andere Personen während dieser Situation fühlen?
Welche Gefühle hast du nach der Situation?

3. Evaluation

Was ist positiv an diesem Verhalten? Was ist negativ daran? Warum?
Gibt dir irgendetwas Zweifel an deinem Verhalten?
Welche Folgen hat dein Verhalten / deine Handlung (für dich oder andere)?
Denke an die Unterrichtsstunden zum entsprechenden Thema – wie beeinflusst dieses Wissen dein Verhalten / deine Handlung und deine Gedanken und Gefühle dabei?
Inwieweit unterstützt du mit deinem Verhalten Nachhaltigkeitsaspekte? (z.B. Umwelt-, wirtschaftliche oder soziale Aspekte)
Wie denkst du jetzt über dein Verhalten?

4. Schlussfolgerung

Denke erneut an die Situation!

Welche anderen Verhaltensweisen / Handlungen wären möglich? Wie bewertest du andere Verhaltens-/ Handlungsmöglichkeiten in Bezug zu deinem Wissen und zu deiner eigenen Einstellung?

Könntest du alternativ handeln und wenn ja, würdest du das wollen? Warum / Warum nicht?

Begründe, ob der Nachhaltigkeitsaspekt in dieser Situation für dich Bedeutung hat!

Könnte und falls ja wie könnte die Situation besser gelöst werden bzw. so, dass alle Beteiligten zufrieden sind?

Welche Fähigkeiten benötigst du, um diese Art von Situation besser lösen zu können?

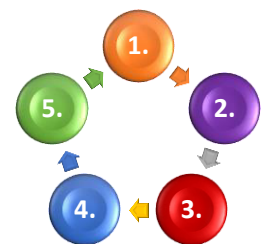
Welche Voraussetzungen müssen getroffen werden, damit man die Situation besser lösen könnte?

5. Plan

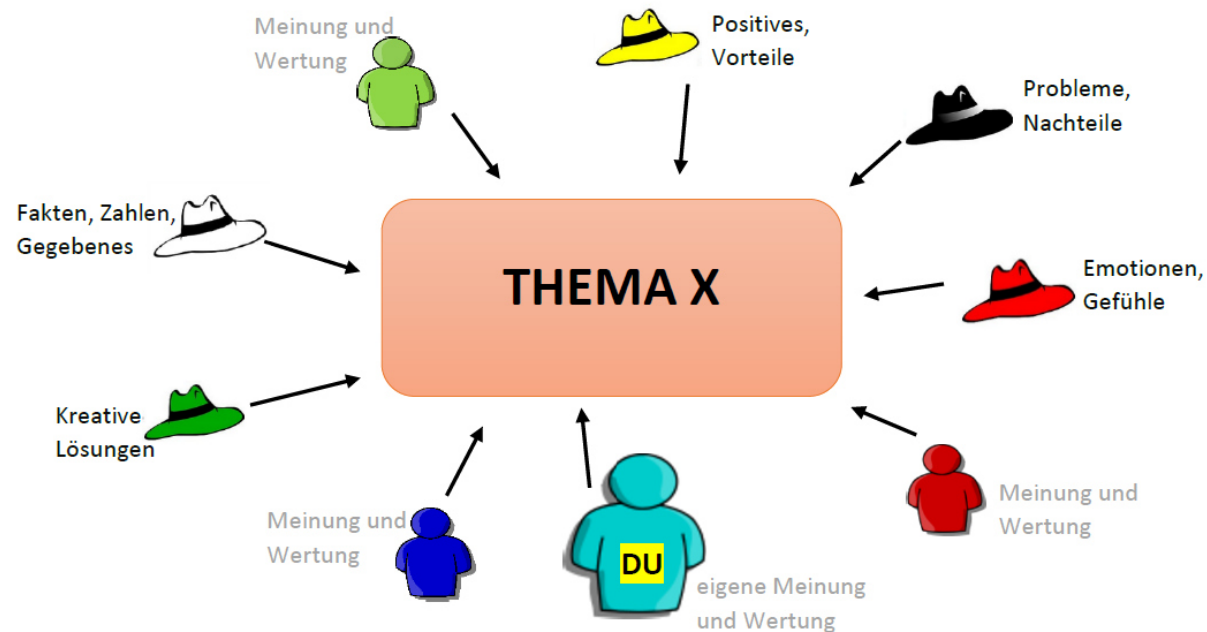
Wenn du wieder in dieser Situation bist, würdest du jetzt etwas anders machen und falls ja, was und warum?

Denke daran: Wenn du dir diese Fragen stellst und sie beantwortest, sollst du immer ehrlich zu dir sein. Du wirst für dein Verhalten nicht kritisiert oder beurteilt, wichtig ist, dass du die Gründe dafür verstehen kannst. Auch wenn du künftig in Situationen nicht anders handeln wirst, ist das nicht falsch und du darfst das natürlich auch ohne schlechtes Gewissen schreiben! Was zählt ist das Nachdenken, warum du etwas so machst oder eben anders und das Herausfinden deiner ganz persönlichen Gründe dafür.

Gerne darfst du Gibbs Reflexivitätszyklus als Hilfsmittel für deine nächsten Blogbeiträge verwenden. Er hilft dir beim zweiten Teil der Aufgabenstellung (Reflektieren über Verhaltensgewohnheit)! Vergiss jedoch nicht, zusätzlich auch über das Thema selbst zu reflektieren!



d. Arbeitsblatt Hütemethode von de Bono

Reflektieren über ein Thema mit der Hüte-Methode

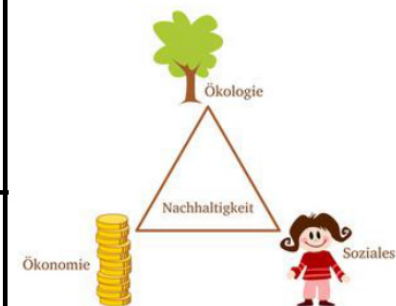
Jeder Person sind die unterschiedlichen Hüte eines Themas mehr oder weniger bewusst. Daraus ergeben sich auf ein Thema unterschiedliche Sichtweisen mit unterschiedlichen Meinungen und Wertungen.

Auch deine Perspektive auf ein Thema ist beeinflusst von dem, was du darüber weißt. Fakten, Gefühle, Vorteile, Nachteile und Lösungsmöglichkeiten, die dir bewusst sind, sind ausschlaggebend dafür, wie du ein Thema betrachtest.

Versuche möglichst viele Hütchen eines Themas aufzudecken und analysiere unterschiedliche Ansichten zu dem Thema. Setze dabei eine kritische Brille bzw. unterschiedliche Hüte auf!

Um unterschiedliche Sichtweisen zu untersuchen und die Hüte aufdecken zu können, orientiere dich am Nachhaltigkeitsdreieck!

Versuche abschließend ein **Fazit** aus deinen Überlegungen abzuleiten!



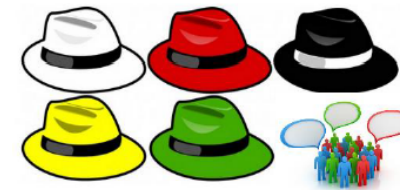
Tipps zum Reflektieren über ein Thema mit der Hüte-Methode

1. Verschaffe dir einen **Überblick über das Thema**. Recherchiere **Fakten**, überlege dir **positive und negative Seiten des Themas, Lösungsmöglichkeiten** und mögliche **Emotionen**, die das Thema bei Personen auslöst (jeweils inkl. Begründungen).
2. Überlege nun, welche **Ansichten und Perspektiven** es auf das Thema geben könnte und versuche, diese anhand deiner Überlegungen von 1. zu begründen. Orientiere dich hierbei am Nachhaltigkeitsdreieck.
3. Beschreibe **deine eigene Haltung** zu dem Thema und begründe sie.
4. Welche **Probleme** ergeben sich aus den unterschiedlichen Sichtweisen? Diskutiere mögliche **Lösungen!**
5. Was ist dein **persönliches Fazit**, nachdem du dich kritisch mit diesem Thema auseinandergesetzt hast?
6. Welches **allgemeine Fazit** kannst du aus deinen Überlegungen ableiten?

Diese Tipps sollen dir bei deinen künftigen Blogbeiträgen helfen!

Wenn du dir unsicher bist, was es heißt, über ein Thema zu reflektieren, erinnere dich an die farbigen Hütchen und lies dir die hier vorgestellten Tipps noch einmal durch und versuche sie in deinem Text umzusetzen!

*Aufgepasst: In **deinen Blogbeiträgen** sollst du immer **sowohl über das Thema als auch über eine deiner Verhaltensgewohnheiten diesbezüglich reflektieren!**
Vergiss bitte keine der beiden Teilaufgaben!*



e. Exemplarischer Blogbeitrag von P18

Ich & Reisen

„Man muss reisen, um zu lernen.“ Ein Zitat von dem berühmtesten amerikanischen Schriftsteller, Mark Twain. Ich glaube, ich bin nicht die einzige Person, die mit ihm einverstanden ist, denn wenn man reist, kann man viel sehen, man kann schöne Landschaften sehen, neue Kulturen, Menschen kennen lernen, Sprachen lernen, sich selbst finden.

Es gibt verschiedene Arten von Reisen. Es gibt Geschäftsreisen, man kann für einen Tag wegfahren oder für mehrere oder auch für immer. Man kann mit unterschiedlichen Mitteln reisen, zum Beispiel mit dem Auto, mit dem Flugzeug oder Schiff. Man kann eine Weltreise machen oder auch in dem eigenen Land eine Reise machen.

Reisen haben jedoch auch ihre Schattenseiten, weil sie meist viel Geld kosten. Heutzutage gibt es immer mehr Menschen, die sich zum Beispiel eine Woche in Italien leisten können, aber es gibt viele, die es nicht tun können. Nehmen wir einen Urlaub in Italien als Beispiel. Wie kommt man nach Italien? Meist mit dem Auto oder Reisebus, vielleicht Flugzeug. Sie haben alle einen hohen CO₂-Ausstoß. Ein anderes Problem ist, dass viele Reisende ihren Müll liegen lassen. Wenn man mit einem Reisebus fährt, kann es sehr unangenehm sein, wenn er überfüllt ist, vor allem im Sommer.

Es gibt für die meisten Probleme eine Lösung. Man soll nachhaltig reisen. Wie macht man das? Lieber mit dem unbequemen Reisebusfahren, als mit dem Auto um den CO₂-Ausstoß zu verringern. Immer darauf achten, dass unser Müll in dem Mülleimer landet. Die Tiere und Pflanzen schützen. Man könnte auch Reiseziele suchen, die in dem eigenen Land sind.

Wenn ich das Wort „Reisen“ höre, muss ich an die lange Fahrt in mein Heimatland denken. Es sind ungefähr 1000 km. Wenn man Glück hat dauert die Fahrt nur 10 Stunden, wenn man im Stau sitzt, fühlt es sich wie eine Ewigkeit an. Egal wie langweilig die Fahrt ist, es lohnt sich, weil nichts schöner ist, als meine Oma zum Lächeln zu bringen.

Die schönste Reise in meinem Leben war unsere Studienfahrt. Wir haben eine Woche in Nizza verbracht. An solche Reisen wird man sich in 50 Jahren immer noch erinnern. Die Alpen persönlich zu sehen ist ganz anders, als auf Bildern. Wenn man einen Film schaut, wird man nie wissen, wie eklig es sich anfühlt, wenn man Meerwasser schluckt. Die hohe Luftfeuchtigkeit fühlt man durch die Kamera auch nicht. Es sind Kleinigkeiten, die man einfach erleben muss. Deswegen werde ich in der Zukunft so viel von der Welt sehen, wie ich kann.

Fazit: Man soll möglichst viel reisen, aber nur nachhaltig.

f. Gegenüberstellung Brendels Modell reflexiven Denkens (2017) und Kategoriensystem zur Bestimmung der Stufen der Reflexionsfähigkeit

Reflexionsstufe				Kriterien zur Bestimmung der Stufe	
Überarbeitung		Ursprüngliche Bezeichnung		Überarbeitung	Ursprüngliche Bezeichnung
K1: Wiedergabe		Stufe 1: Wiedergabe		<ul style="list-style-type: none"> - Reine Reproduktion von Inhalten oder Gegebenheiten - <u>Faktenwiedergabe</u> - <u>Sachliche Fragen, die anschließend beantwortet werden</u> 	Reine Reproduktion von Inhalten oder Gegebenheiten (keine Reflexion)
K2: Bezug nehmen		Stufe 2: Bezug nehmen		<p>Die Schülerin oder der Schüler:</p> <ul style="list-style-type: none"> - äußert eine Beobachtung, Meinung oder Wertung, jedoch ohne Begründung - stellt rhetorische Fragen ohne Antwort - drückt Gefühle aus 	<p>Die Schülerin oder der Schüler:</p> <ul style="list-style-type: none"> - transformiert oder konzeptualisiert die Unterrichtsinhalte geringfügig - äußert eine Beobachtung, Meinung oder Wertung, jedoch ohne Begründung oder Folgerung - stellt rhetorische Fragen ohne Antwort - drückt Gefühle aus - nennt Lernzuwachs
K3: <u>Begründung/ Bewertung</u>	3a: Begründung		K3a: Begründung	<p>Die Schülerin oder der Schüler:</p> <ul style="list-style-type: none"> - begründet Prozesse oder Sachverhalte oberflächlich, Handlungsbedarf wird erkannt - <u>Urteil mit simpler Begründung</u> 	<p>Die Schülerin oder der Schüler:</p> <ul style="list-style-type: none"> - begründet Prozesse oder Sachverhalte oberflächlich, Handlungsbedarf wird erkannt - Urteil mit simpler Begründung

	3b: <u>Be- wertung</u>	Stufe 3: Zusam- men- hänge	K3b: Ver- netzung	<p>Die Schülerin oder der Schüler:</p> <ul style="list-style-type: none"> - stellt persönlichen Zusammenhang zu den Inhalten her, z.B. durch Verbindung mit Vorwissen oder eigenen Erfahrungen - versucht oberflächlich Wechselbeziehungen darzustellen - kommt zu Fazit ohne weiter auf vorher genannte Punkte einzugehen 	<p>Die Schülerin oder der Schüler:</p> <ul style="list-style-type: none"> - stellt persönlichen Zusammenhang zu den Inhalten her, z.B. durch Verbindung mit Vorwissen oder eigenen Erfahrungen - versucht oberflächlich Wechselbeziehungen darzustellen
<u>K4: Analyse</u>		Stufe 4: Beurteilung		<p>Voraussetzung: Konzeptualisierung auf gehobenem Niveau</p> <p>Die Schülerin oder der Schüler:</p> <ul style="list-style-type: none"> - verknüpft Inhalte mit z. B. Theorien oder persönlichen Erfahrungen - führt vertiefte Begründungen für Prozesse, Sachverhalte oder fremde Handlungen an - analysiert ein Problem, sucht nach Antworten und Alternativen, spekuliert oder stellt Hypothesen auf - untersucht den Zusammenhang zwischen Theorie und Praxis - erkennt die Komplexität des Sachverhalts - <u>untersucht Thema aus unterschiedlichen Perspektiven</u> 	<p>Voraussetzung: Konzeptualisierung auf gehobenem Niveau</p> <p>Die Schülerin oder der Schüler:</p> <ul style="list-style-type: none"> - verknüpft Inhalte mit z. B. Theorien oder persönlichen Erfahrungen - führt vertiefte Begründungen für Prozesse, Sachverhalte oder eigene und fremde Handlungen an - analysiert ein Problem, sucht nach Antworten und Alternativen, spekuliert oder stellt Hypothesen auf - untersucht den Zusammenhang zwischen Theorie und Praxis - erkennt die Komplexität des Sachverhalts
K5: (Re)Konstruktion		Stufe 5: (Re)Konstruktion		<p>Voraussetzung: Konzeptualisierung auf hohem Abstraktionsniveau</p> <p>Die Schülerin oder der Schüler:</p> <ul style="list-style-type: none"> - kommt zu persönlichem Fazit und bezieht systematisch und differenziert Stellung 	<p>Voraussetzung: Konzeptualisierung auf hohem Abstraktionsniveau</p> <p>Die Schülerin oder der Schüler:</p> <ul style="list-style-type: none"> - kommt zu persönlichem Fazit oder bezieht systematisch und differenziert Stellung

		(vorherige Auseinandersetzung mit unterschiedlichen Perspektiven nötig) - formuliert eine persönliche Theorie oder leitet Generalisierungen ab	- formuliert eine persönliche Theorie oder leitet Generalisierungen ab - erweitert und modifiziert die Unterrichtsergebnisse deutlich über das angestrebte Niveau hinaus
--	--	---	---

Gegenüberstellung Zimmermann und Welzels Stufenmodell (2008) und Kategoriensystem zur Bestimmung der Stufen der Reflexivitätsfähigkeit

Reflexivitätsstufe			Definition		
Überarbeitung		Ursprüngliche Bezeichnung	Überarbeitung		Ursprüngliche Bezeichnung
K1: Sachbezogene Beschreibung	K1a: Objektive, sachbezogene Beschreibung	Stufe 1: Sachbezogene Beschreibung	<ul style="list-style-type: none"> - keine explizite Reflexion - Einfache Wiedergabe von Handlungsabläufen 	<ul style="list-style-type: none"> - <u>Verallgemeinerung der persönlichen Sicht ohne Begründungen</u> 	<ul style="list-style-type: none"> - bezogen auf konkretes Handeln - einfache Wiedergabe von Handlungsabläufen - Beschreibung von Situationen ohne zusätzliche Kommentare, ohne eigene Gedanken - Erste Distanzierung vom konkreten Tun ohne Begründung, keine explizite Reflexion
	K1b: Subjektive, sachbezogene Beschreibung			<ul style="list-style-type: none"> - Beschreibung von Situationen <u>aus der Ich-Perspektive</u> ohne zusätzliche Kommentare, ohne eigene Gedanken, <u>ohne Begründung, ohne Rechtfertigung, ohne Bewertungen</u> 	
K2: Handlungsbezogene Begründung	K2a: Handlungsbezogene Begründung für allgemeine Handlung	Stufe 2: Handlungsbezogene Begründung	<ul style="list-style-type: none"> - Beschreibung von Sachverhalten inklusive Begründungen - <u>Bewertungen von allgemeinen Handlungen mit Begründungen</u> - <u>Konkrete Handlungsaufforderung</u> - Nennung von allgemeinen Handlungsvorschlägen/ -alternativen 		<ul style="list-style-type: none"> - Beschreibung von Sachverhalten - Sinnggebung mit Begründungen oder Handlungsalternativen werden aufgeführt - Konkrete Handlungsebene

	K2b: Handlungsbezogene Begründung für eigene Handlung		<ul style="list-style-type: none"> - <u>Beschreibung von persönlichen Handlungen mit zusätzlichen Kommentaren, eigenen Gedanken, Begründungen, Rechtfertigungen oder Bewertungen</u> - konkrete Handlungsebene (<u>d.h. keine Verallgemeinerung</u>) - innere Distanzierung vom bloßen Handeln, da Formulierungen der eigenen Perspektive vorhanden - Nennung von persönlichen Handlungsvorschlägen/ -alternativen - <u>künftige Handlungsplanungen ohne Gründe dafür zu nennen</u> 	<ul style="list-style-type: none"> - Innere Distanzierung vom bloßen Handeln, da Formulierungen der eigenen Perspektive - vorhanden
K3: Analytische Abstraktion	K3a: analytische Abstraktion für allgemeine Handlung	Stufe 3: Analytische Abstraktion	<p>Sowohl für allgemeine als auch persönliche Handlungen:</p> <ul style="list-style-type: none"> -gemachte Erfahrungen werden unter Berücksichtigung mehrerer Perspektiven/<u>von einer höheren Perspektive aus</u> analysiert -<u>Abwägung und Beurteilung unterschiedlicher Ansätze</u> - ein neues, erweitertes Verständnis wird entwickelt (<u>Fazit aus Überlegungen</u>) - <u>Dialog mit der Situation und mit sich selbst</u> - aus der Analyse heraus werden Konsequenzen im Hinblick auf zukünftiges Verhalten, zukünftige Aktivitäten gezogen 	<ul style="list-style-type: none"> - Gemachte Erfahrungen werden unter Berücksichtigung mehrerer Perspektiven (Personen und Geleertes) analysiert - Elemente und Strukturen werden auf einer allgemeinen Ebene erkannt, identifiziert und zueinander in Beziehung gesetzt - aus dieser Analyse heraus werden Konsequenzen im Hinblick auf zukünftiges Verhalten, zukünftige Aktivitäten gezogen - ein neues, erweitertes Verständnis wird entwickelt
	K3b: analytische Abstraktion für eigene Handlung			
K4: Kritischer Diskurs		Stufe 4: Kritischer Diskurs	<ul style="list-style-type: none"> - die Erkenntnisse werden im Hinblick auf übergeordnete Werte, Prinzipien und/oder auf relevante Theorien hin bewertet - es erfolgt eine kritische Auseinandersetzung damit (<u>Multiperspektivität Voraussetzung dafür</u>) 	<ul style="list-style-type: none"> - die Erkenntnisse werden im Hinblick auf übergeordnete Werte, Prinzipien und/oder auf relevante Theorien hin bewertet - es erfolgt eine kritische Auseinandersetzung damit

		<ul style="list-style-type: none"> - Zunahme der Fähigkeit zum Erkennen eigener Anteile und Deutungen in Interaktionen; <u>Bewusstsein, dass frühere Erfahrungen das eigene Handeln beeinflussen; Bewusstsein über eigene Grenzen</u> - <u>Bewusstsein, dass Handeln eingebettet ist in verschiedene Kontexte (z.B. politisch, historisch, gesellschaftlich), die Einfluss nehmen</u> 	<ul style="list-style-type: none"> - Zunahme der Fähigkeit zum Erkennen eigener Anteile und Deutungen in Interaktionen
K 5: Professionalisierung	Stufe 5: Professionalisierung	<ul style="list-style-type: none"> - Erkennen des eigenen Urteils-Handlungs-Hiatus UND Durchdenken, Planen von zielorientierten Verbesserungsstrategien - konkrete Verbesserungsmaßnahmen (am eigenen Tun) werden initiiert - festgefahrene Rollenverständnisse werden thematisiert und konstruktiv in Frage gestellt, andere Handlungsmöglichkeiten diesbezüglich ausgelotet und ausprobiert 	<ul style="list-style-type: none"> - Erkennen des eigenen Urteils-Handlungs-Hiatus UND Durchdenken, Planen von zielorientierten Verbesserungsstrategien - konkrete Verbesserungsmaßnahmen (am eigenen Tun, an professionellem Umfeld) werden initiiert - festgefahrene Rollenverständnisse werden thematisiert und konstruktiv in Frage gestellt, andere Handlungsmöglichkeiten diesbezüglich ausgelotet und ausprobiert

g. Einzelauswertungen im Rahmen der qualitativen Inhaltsanalyse

Testperson P1

Allgemeines zur Person

Name	P1
Klasse	1
Geschlecht	Männlich
Anzahl Beiträge	10 Beiträge
Reihenfolge Reflektiermethoden	Hüte – Gibbs
Verständnis von Reflektieren (vor Blogprojekt)	Das Auseinandersetzen mit gesellschaftlichen, politischen und ökologischen Sachverhalten.
Verständnis von Reflektieren (nach Blogprojekt)	Reflektieren ist das kritische Hinterfragen alltäglicher Dinge. Es ist eine Einstellung, die skeptisch aber auch zielgerichtet und konstruktiv ist und somit die Werte der Aufklärung widerspiegelt. Ähnlich wie bei der Philosophie soll das eigene Denken, dessen Wurzeln sowie Konsequenzen desselben bewusst wahrgenommen werden, um tieferes Verständnis zu erlangen.
Gibbs Reflexivitätszyklus Verständlichkeit	Stimmt voll und ganz
Gibbs Reflexivitätszyklus Hilfestellung	Stimmt eher schon
Hütemethode Verständlichkeit	Stimmt voll und ganz
Hütemethode Hilfestellung	Stimmt eher schon

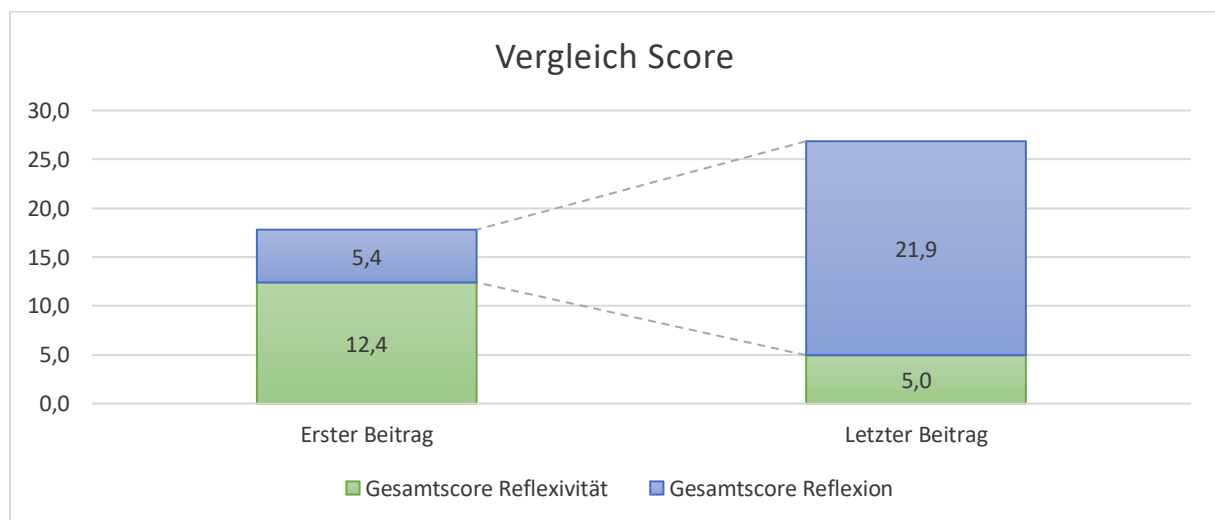
P1 ist männlich und in Klasse 1. Insgesamt gibt es zehn Beiträge von P1, nach den ersten drei Beiträgen erhält die Testperson die Hütemethode, anschließend Gibbs Reflexivitätszyklus. Zu Beginn des Projekts versteht P1 folgendes unter dem Begriff *Reflektieren*: „Das Auseinandersetzen mit gesellschaftlichen, politischen und ökologischen Sachverhalten“ (P1). Am Ende des Projekts formuliert er folgende Definition: „Reflektieren ist das kritische Hinterfragen alltäglicher Dinge. Es ist eine Einstellung, die skeptisch aber auch zielgerichtet und konstruktiv ist und somit die Werte der Aufklärung widerspiegelt. Ähnlich wie bei der Philosophie soll das eigene Denken, dessen Wurzeln sowie Konsequenzen desselben bewusst wahrgenommen

werden, um tieferes Verständnis zu erlangen“ (P1). Die erste Definition ist noch etwas ungenau und schließt die Aspekte der Gesellschaft, Politik und Ökologie ein. Die letzte Definition ist wesentlich ausführlicher und konzentriert sich vor allem auf die Fähigkeit der Reflexion. Einstellungen sollen analysiert werden und Gründe für das eigene Denken aufgedeckt werden. Das Reflektieren in der letzten Definition kann alltägliche Dinge betreffen und ist somit weiter ausgelegt als das Reflektieren der ersten Definition. Beide Reflektierhilfen sind für P1 *voll und ganz* verständlich und *eher schon* hilfreich für das Verfassen der eigenen Blogbeiträge.

Vergleich der Beiträge zum Thema Tierschutz

Score

P1 hat beim ersten Beitrag einen Gesamtscore im Reflektieren von 17,8, der sich auf einen Reflexionsscore von 5,4 und einen Reflexivitätsscore von 12,4 verteilt. Der zweite Beitrag zum Thema Tierschutz weist einen Gesamtscore im Reflektieren von 26,9 auf und ist somit um 9,1 höher als der Score des ersten Beitrags. Der Reflexivitätsscore beträgt im zweiten Beitrag 5,0, der Reflexionsscore 21,9. Folglich steigt der Reflexionsscore um den Wert 16,5, während der Reflexivitätsscore um den Wert 7,4 sinkt.



Analyse der Reflektierstufen

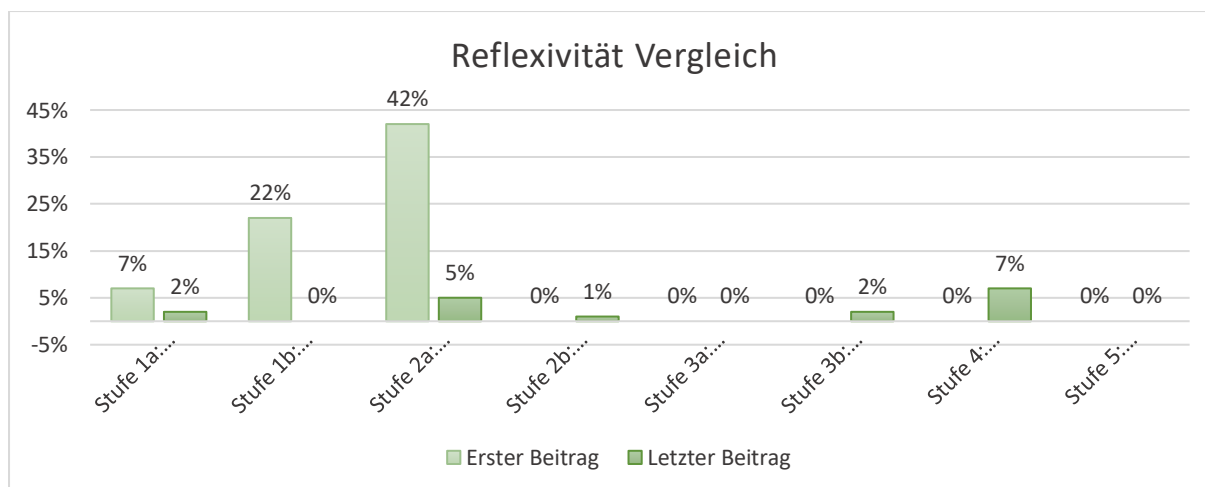
Der erste Beitrag zum Thema Tierschutz von P1 besteht zu 71% aus Reflexivitätsanteilen. 7% des Gesamttextes weisen die Reflexivitätsstufe 1a, 22% die Reflexivitätsstufe 1b und 42% die Reflexivitätsstufe 2a auf. Die Testperson begründet überwiegend allgemeines Verhalten zum Thema Tierschutz, sein eigenes Verhalten beschreibt er lediglich ohne Gründe dafür zu nennen. So schreibt er beispielsweise über sein eigenes Verhalten: „Kaum jemand engagiert sich oder ändert sein Konsumverhalten - mich eingeschlossen“ (P1, erster Beitrag, Reflexivitätsstufe 1b). Zum allgemeinen Verhalten formuliert er folgende Sätze:

Entscheidend für die meisten Konsumenten [...] ist der Preis bzw. der Genuss. Moral, Ethik sowie Nachhaltigkeit werden nur am Rande berücksichtigt. Erziehung und internalisierte Verhaltensweisen tun auch ihren Teil. (P1, erster Beitrag, Reflexivitätsstufe 2a)

Im letzten Beitrag, der zu 17% aus Reflexivitätsanteilen besteht, gehören 2% des Textes zu Stufe 1a, 5% zu Stufe 2a, 1% zu Stufe 2b, 2% zu Stufe 3b und 7% zu Stufe 4. P1 erreicht im zweiten Beitrag zum Thema Tierschutz maximal Stufe 4 und somit zwei Reflexivitätsstufen höher als im ersten Beitrag. Er beschreibt und begründet allgemeines Verhalten und sucht nach möglichen Handlungsalternativen. Folgender Textauszug entspricht beispielsweise Stufe 2a im letzten Beitrag: „Der Preis an der Ladentheke und der Genuss scheinen unser Verhalten stärker zu beeinflussen, als es abstrakte Gedanken über den Ursprung der Menschenwürde tun“ (P1, letzter Beitrag, Reflexivitätsstufe 2a). Über sein persönliches Verhalten reflektiert er auf hohem Niveau und erreicht dabei einen kritischen Diskurs:

Ich fühle mich heuchlerisch, wenn ich hier ausschweifend über die Relevanz von Tierschutz schreibe und gleichzeitig Fleisch esse, lieber mit dem Auto, als mit dem Rad fahre und auch nicht vorhabe, dies alles zeitnah zu ändern. Ich muss wohl abwägen, Kompromisse machen und einsehen, dass ich die Welt nicht sofort ändern kann. Zumal ich auch denke, dass es momentan dringlichere Probleme gibt. Es fühlt sich für mich falsch an, den Schutz von Mastschweinen zu fordern, wenn es gleichzeitig Menschen gibt, denen es ähnlich schlecht oder sogar schlechter geht. Vielleicht suche ich aber auch nur nach Ausreden, um nicht aktiv werden zu müssen. Die Veränderung unseres Verhaltens verläuft graduell durch Sensibilisierung, Aufklärung, und das Fühlbarmachen von kalten Zahlen. (P1, letzter Beitrag, Reflexivitätsstufe 4)

Obwohl der Gesamtreflexivitätsanteil am Text vom ersten zum letzten Beitrag abnimmt, steigt das gezeigte Reflexivitätsniveau deutlich an und P1 beschreibt und begründet Handlungen nicht nur, sondern analysiert allgemeines und sein persönliches Verhalten tiefgründiger.



P1 reflektiert über das Thema Tierschutz lediglich in der Reflexionsstufe 2 und äußert seine Beobachtung und Meinung diesbezüglich, jedoch ohne weitere Analyse. Insgesamt können

27% des ersten Textes der Stufe 2 der Reflexionsfähigkeit zugeordnet werden. Die beiden folgenden Sätze sind Beispiele daraus: „Bilder von unwürdigen Legebatterien und von Ferkeln, die auf wenigen Quadratmetern hausen, sind keinem von uns fremd. Überall in den Medien wird darüber aufgeklärt, diskutiert und zum Nachdenken angeregt“ (P1, erster Beitrag, Reflexionsstufe 2). Der letzte Beitrag weist einen höheren Textanteil mit Reflexionsstufen auf. Insgesamt wird in 83% des Textes über das Thema reflektiert, wobei 15% der Stufe 1, 33% der Stufe 2, 6% der Stufe 3a, 8% der Stufe 3b, 13% der Stufe 4 und 8% der Stufe 5 entsprechen. Ähnlich wie beim ersten Beitrag reflektiert P1 überwiegend in Form von der Nennung der eigenen Meinung und in der Beschreibung von eigenen Beobachtungen. Darüber hinaus nennt er nun jedoch auch Begründungen und beurteilt Sachverhalte aus unterschiedlichen Perspektiven. Die Testperson erreicht ein sehr hohes Abstraktionsniveau, das Voraussetzung für Stufe 5 ist, die die Testperson ebenfalls erreicht und somit alle Reflexionsstufen im letzten Beitrag von P1 enthalten sind. Folgender Textausschnitt zeigt P1 hohes und sehr hohes Abstraktionsniveau im letzten Beitrag, das im ersten Beitrag nicht vorhanden ist. Zunächst wird das Problem der Grenzziehung untersucht, anschließend äußert er seine persönliche Einstellung dazu und nimmt differenziert Stellung:

Es ist absurd. Wir legitimieren unser amoralisches Verhalten gegenüber Tieren, indem wir ihr Unvermögen zu moralischen Empfindungen unterstreichen. In unserem Versuch uns abzugrenzen, machen wir uns aber gleich mit Tieren.

Ein riesiges Problem in der tierethischen Debatte ist die Grenzziehung, für welche man nie vollkommen objektive Maßstäbe finden kann. Schnell kann man sich darauf einigen, dass Schimpansen besonderen Schutz verdient haben. Aber gilt dies auch für Koalas, Waschbären, Ratten oder gar Insekten? Schon in der griechischen Antike wusste man, dass Moral nie exakt sein kann. Das heißt aber nicht, dass da wo die Grenzen nicht klar gezogen werden können, keine Grenzen existieren. Unsere Aufgabe als Gesellschaft ist es, durch den ständigen Diskurs die Grenzen immer wieder neu zu definieren.

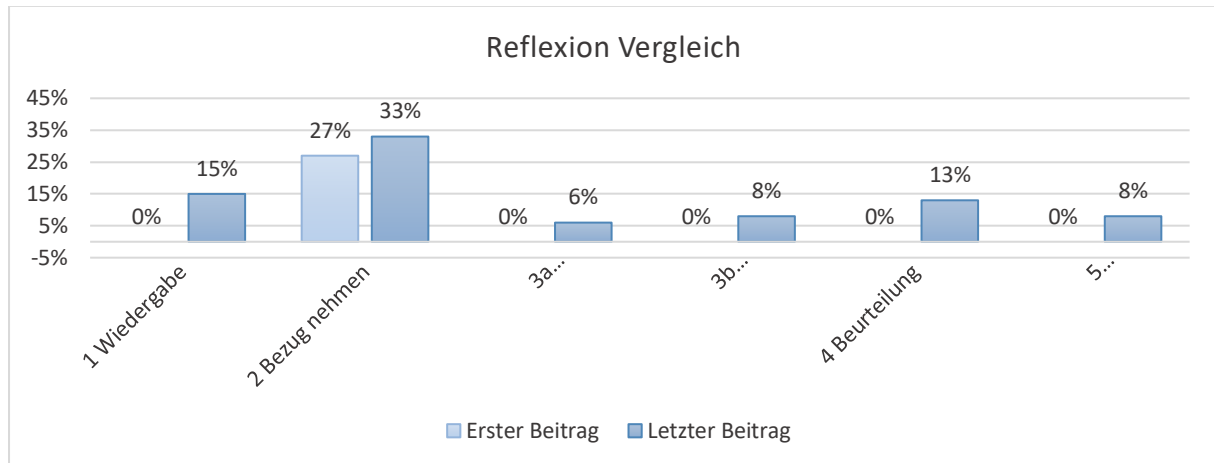
Wir müssen abwägen: Wenn ein hohes menschliches Interesse einem relativ geringen tierischen Interesse gegenübersteht, ist es zu billigen. In Bezug auf die Jagd heißt das, dass die Treibjagd aus Spaß am Jagen verboten sein muss, das Jagen von Wildschweinen zur Regulation des Bestands hingegen akzeptabel ist.

Die Aussicht auf die Heilung von Parkinson für Millionen von Menschen rechtfertigt das Leid von tausenden Labormäusen (nach utilitaristischer Auffassung), aber unser Wunsch nach Schönheit kann nicht das Leid von Schimpansen für Kosmetikstudien aufwiegen. Gutes Aussehen ist kein dringendes Lebensinteresse des Menschen und darf daher nicht als solches behandelt werden. Leider ist es oft genau umgekehrt. Ein geringes menschliches Interesse (Spaß an Gänseleberpastete) rechtfertigt ein gewaltiges Interesse des Tieres (Ausbleiben von großen Qualen).

(P1, letzter Beitrag, zunächst Reflexionsstufe 4, anschließend Stufe 5)

Die Reflexionsfähigkeit von P1 wird im letzten Beitrag nicht nur häufiger gezeigt, die Testperson erreicht auch höhere Reflexionsstufen und generalisiert auf einem gehobenen Niveau.

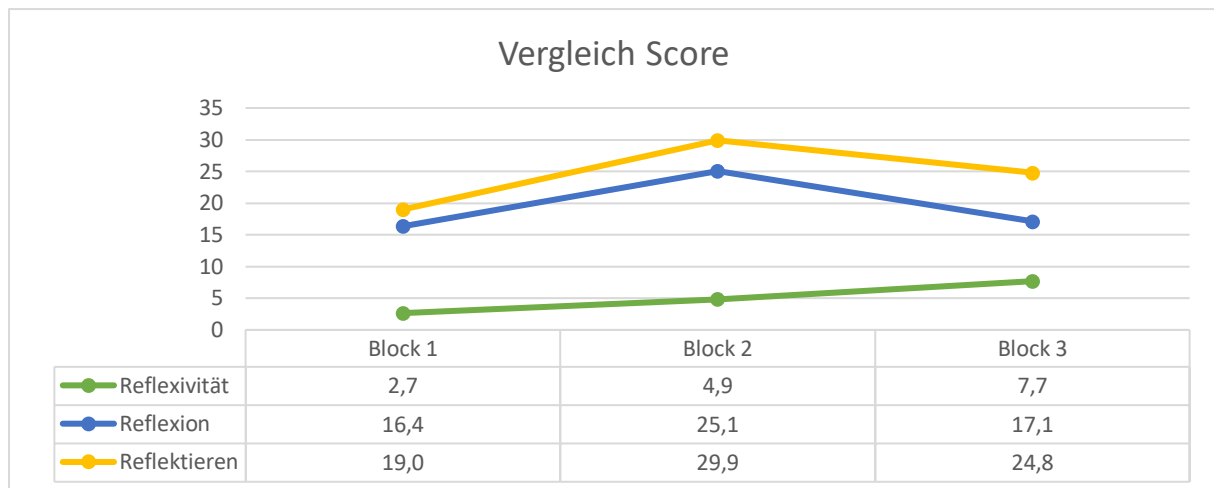
Folglich ist sowohl die Reflexivitätsfähigkeit als auch die Reflexionsfähigkeit in Bezug auf die Analyse der einzelnen Stufen im letzten Beitrag von P1 höher als in seinem ersten Beitrag.



Vergleich der Beiträge vor, zwischen und nach den Reflektiermethoden

Score

Die Beiträge vor der ersten Reflektierhilfe weisen bei P1 im Mittel einen Gesamtscore im Reflektieren von 19 auf, der sich aus einem Reflexionsscore von 16,4 und einem Reflexivitätsscore von 2,7 zusammensetzt. Nach Erhalt der ersten Reflektierhilfe, in diesem Fall der Hütemethode, steigt der Gesamtscore um den Wert 10,9 auf 29,9 und erreicht dabei seinen Höchstwert. Der Wert untergliedert sich in einen Reflexivitätsscore von 4,9 und einen Reflexionsscore von 25,1. Beide Scores sind höher als die Scores vor der Hütemethode. Die Beiträge nach der zweiten Reflektierhilfe – bei P1 Gibbs Reflexivitätszyklus – haben einen durchschnittlichen Gesamtscore von 24,8, der um 5,8 höher ist als der Score vor den beiden Reflektierhilfen und um 5,1 niedriger als der Score nach Erhalt der ersten Reflektierhilfe. Der Reflexionsscore im letzten Textblock beträgt 17,1 und der Reflexivitätsscore 7,7. Während die Reflexionsfähigkeit in Hinblick auf den Score zunächst von Textblock 1 auf Textblock 2 steigt (mit Erhalt der Hütemethode), fällt sie anschließend auf einen Wert knapp über den Ausgangsscore ab. Die Reflexivitätsfähigkeit hingegen steigt kontinuierlich an.



Analyse der Reflektierstufen

P1 reflektiert in den ersten Beiträgen vor Erhalt einer Reflektierhilfe über das Verhalten in den Reflexivitätsstufen 1b, 2a und 2b mit 3%, 5% und 6% Anteil am Gesamttext. Somit enthalten die Texte durchschnittlich nur einen geringen Reflexivitätsanteil von 14%. Er reflektiert beispielsweise über sein eigenes Handeln und begründet in Hinblick auf das Thema Fortbewegungsmittel seine Entscheidung für sein gewähltes Fahrzeug. Außerdem begründet er, warum allgemein Autos so häufig genutzt werden. Auch nach Erhalt der ersten Reflektierhilfe (Hüttemethode) ändert sich seine Reflexivitätsfähigkeit in Hinblick auf die gezeigten Reflexivitätsstufen nur gering. P1 zeigt nun Stufe 1a und nicht mehr Stufe 1b und reflektiert häufiger über allgemeines Verhalten inklusive Begründungen und Handlungsalternativen. Ein Beispiel hierfür ist im fünften Beitrag zu finden:

Alternativen gibt es schließlich zu Genüge, auch wenn sie vielleicht nicht ganz so lecker sind. Regionales und saisonales Gemüse. Dabei muss natürlich nicht ganz auf Fleisch verzichtet werden. Eine Mäßigung in Verbindung mit Wertschätzung gegenüber Fleisch wäre zumindest ein guter Anfang. (P1, fünfter Beitrag, Reflexivitätsstufe 2a)

Im Durchschnitt kommen Stufe 1a in 5%, Stufe 2a in 15%, Stufe 2b in 5% und Stufe 3b in 1% des Gesamttextes vor und der Gesamttext besteht zu 26% aus Reflexivitätsanteilen. Stufe 3b wird nur in einem Beitrag erreicht, in dem P1 als Fazit einen Vorsatz für sich selbst schildert: „Mäßigung als Lebensmotto. Mein neuer Vorsatz für das Jahr 2018“ (P1, fünfter Beitrag, Reflexivitätsstufe 3b). Die Beiträge nach Erhalt der zweiten Reflektierhilfe (Gibbs) bestehen mit Ausnahme eines Beitrags aus Textanteilen, die der Reflexivitätsstufe 4 zugeordnet werden können. Hier analysiert P1 so tiefgründig, dass ein kritischer Diskurs sichtbar wird. In den Beiträgen von Block 3 kommen alle Reflexivitätsstufen vor und machen durchschnittlich 26% des Gesamttextes aus (4% Stufe 1a, 1% Stufe 1b, 4% Stufe 2a, 1% Stufe 2b, 1% Stufe 3a, 9% Stufe 3b und 6% Stufe 4), außer der höchsten Stufe 5. Die Analyse und das Ableiten von

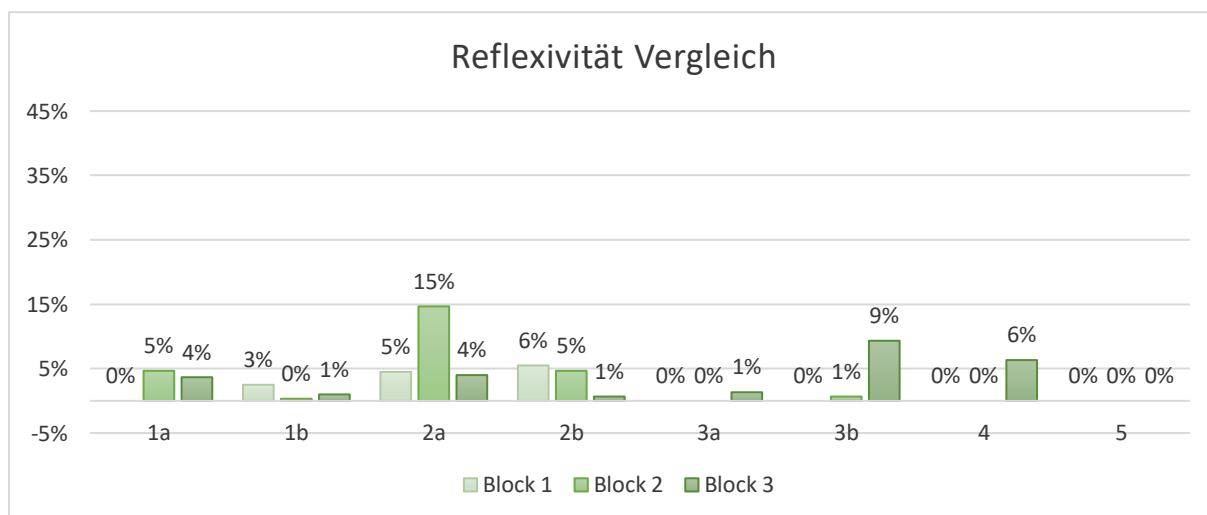
künftigen Handlungen bezogen auf das eigene Verhalten kommen am häufigsten vor, wie beispielsweise im folgenden Textausschnitt:

Die Bilder von durstenden Kindern lassen trotzdem niemanden kalt, mich auch nicht. Dennoch bin ich skeptisch, ob ich wirklich etwas bewegen kann. Patentrezepte gibt es schließlich keine. Trotzdem werde ich versuchen meinen Fleischkonsum Schritt für Schritt zu reduzieren. Auch Hosen und Baumwollshirts werde ich künftig mit Bedacht kaufen. Wenn ich mich bei jedem Einkauf frage, ob es wirklich sein muss und welche Folgen der Kauf für Umwelt und Natur hat, dann ist schon viel getan. (P1, siebter Beitrag, Reflexivitätsstufe 3b)

Beim kritischen Diskurs werden P1 seine persönlichen Grenzen bewusst und er untersucht, wodurch sein Handeln alles beeinflusst wird:

Na ja, damit werde ich die Welt nicht einreißen können, was ernüchternd für mich ist. Am guten Willen wird es schließlich nicht scheitern. Das Thema erschlägt einen und ich habe oft das Gefühl, dass die intellektuelle Verwirrung nicht geringer wird, je länger ich mich mit einem speziellen Thema beschäftige. Oft liest man Dinge, die nicht so recht ins eigene Weltbild passen. Wenn ich dann den ersten Impuls des Verdrängens und Abtuens überwunden habe, folgt meistens Verwirrung, aus der wiederum Resignation werden kann. (P1, neunter Beitrag, Reflexivitätsstufe 4)

P1 entwickelt sich in Hinblick auf die einzelnen Reflexivitätsstufen. Zunächst reflektiert er eher oberflächlich, mit Erhalt von Gibbs Reflexivitätszyklus reflektiert er wesentlich tiefgründiger und abstrahiert auf einem hohen Niveau. Die maximal erreichte Reflexivitätsstufe steigt von 2b in Block 1 auf 3b in Block 2 (wobei dies nur einmal erreicht wird) auf 4 in Block 3, die durchschnittlich 6% des Gesamttextes ausmacht.

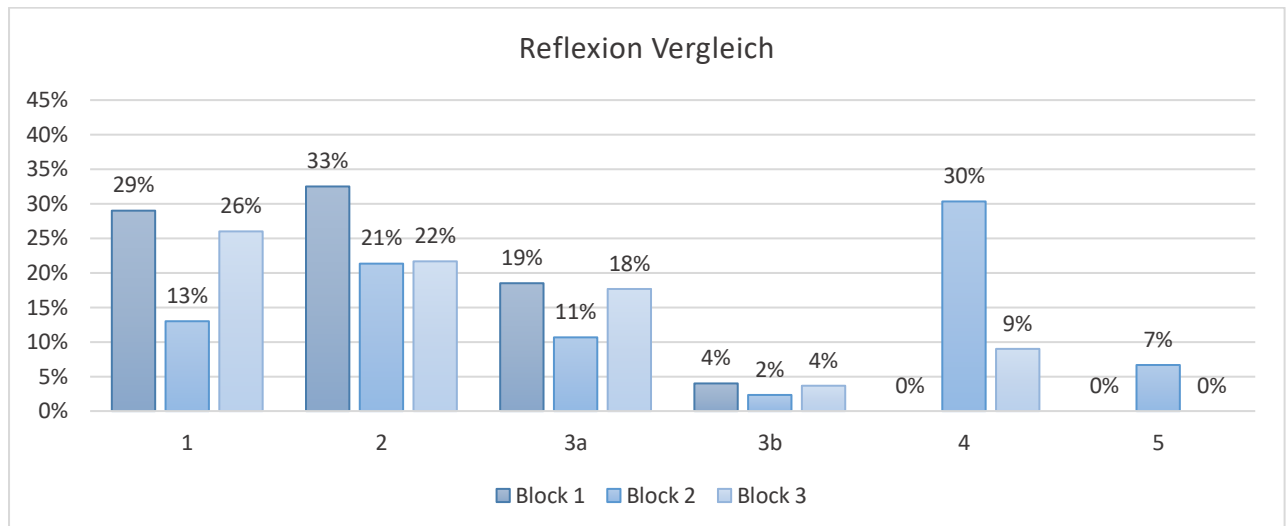


Die Beiträge von P1 im ersten Textblock und somit vor Erhalt der ersten Reflektierhilfe haben einen durchschnittlichen Gesamtreflexionsanteil von 85% und bestehen aus den Stufen 1, 2, 3a und 3b (mit 29%, 33%, 19% und 4% Anteil am Text). Die Testperson reflektiert über das Thema überwiegend in Form von Beschreibungen und Wertungen jedoch ohne Begründungen, wie z.B. „Schaut man auf deutsche Straßen, so ist man schnell ernüchtert. Der aktuelle

Trend geht hin zum großen, spritfressenden Geländewagen - vor allem in Großstädten“ (P1, zweiter Beitrag, Reflexionsstufe 2). Die Stufe 3b kommt nur einmal als Fazit aus den Überlegungen vor, ansonsten bleiben die Begründungen allgemeiner Art. Nach Erhalt der Hütemethode haben die Beiträge durchschnittlich einen Gesamtreflexionsanteil von 84%, der bis auf ein Prozent mit dem vorherigen Gesamtanteil am Text übereinstimmt. Als höchste Stufe wird nun Stufe 5 erreicht, die eine Reflexion auf einem sehr hohen Generalisierungsniveau beinhaltet. Folgender Abschnitt enthält die Kriterien für diese Stufe:

Stetiger Fortschritt treibt uns um. Rüdiger Safranski sagte dazu, dass damals derjenige, der Fortschritte haben wollte, sich erklären musste, wohingegen das Festhalten am Status Quo die Regel war - politisch wie in Bezug auf Konsum. Nun sind wir aber in den westlichen Nationen an einem Punkt angekommen, an dem Wachstum absurd ist. Wir haben doch von allem zu viel. Vielmehr müssen wir an qualitativen Wachstum statt des quantitativen arbeiten. Vor allem aber müssen wir an eine gerechte Verteilung der Güter denken. Ein Lebensstil wie wir ihn haben, würde uns binnen eines Jahres in Müll und Abgasen versinken lassen, wenn er global geführt werden würde. (P1, vierter Beitrag, Reflexionsstufe 5)

P1 kommt zu einem persönlichen Fazit und bezieht systematisch und differenziert Stellung, nachdem er sich zuvor mit unterschiedlichen Perspektiven auseinandergesetzt hat. Die Stufe 4, bei der in Form von vertieften Begründungen und Problemanalysen reflektiert wird, wird im zweiten Textblock mit 30% am häufigsten gezeigt. Nach Erhalt der zweiten Reflektierhilfe (Gibbs Reflexivitätszyklus) erreicht P1 maximal Stufe 4 mit durchschnittlich 9% Anteil am Gesamttext pro Beitrag. Die Stufen 1 bis 3b sind ebenfalls alle vertreten mit 26%, 22%, 18% und 4% Anteil am Text. Insgesamt enthalten im dritten Textblock 79% der Textanteile Reflexionsstufen. Die Hütemethode, die P1 als erste Hilfe erhält, hat einen deutlichen Einfluss auf die Reflexionsstufenverteilung, die sich nach oben hin verschiebt. Die Testperson reflektiert darauf hin auf einem deutlich höheren Niveau, während die zweite Reflektierhilfe keinen sichtbaren positiven Einfluss mehr nimmt. Der prozentuale Anteil beider Reflektierfähigkeiten an den einzelnen Texten bleibt in den drei Textblöcken nahezu gleich, sodass die Reflektierhilfen darauf keinen Einfluss nehmen.



Testperson P2

Allgemeines zur Person

Name	P2
Klasse	1
Geschlecht	Männlich
Anzahl Beiträge	10 Beiträge
Reihenfolge Reflektiermethoden	Gibbs - Hüte
Verständnis von Reflektieren (vor Blogprojekt)	über etwas nachzudenken oder etwas wiederzugeben im Sinne von "widerspiegeln"
Verständnis von Reflektieren (nach Blogprojekt)	Reflektieren bedeutet für mich, über eine ausgewählte Thematik genau und aus verschiedenen Perspektiven nachzudenken. Zu Beginn steht dabei immer eine Analyse der momentanen Situation, gefolgt von den möglichen Handlungsoptionen und schließlich einem Fazit in dem die zukünftigen Vorgehen ersichtlich werden.
Gibbs Reflexivitätszyklus Verständlichkeit	Stimmt eher schon
Gibbs Reflexivitätszyklus Hilfestellung	Stimmt eher nicht

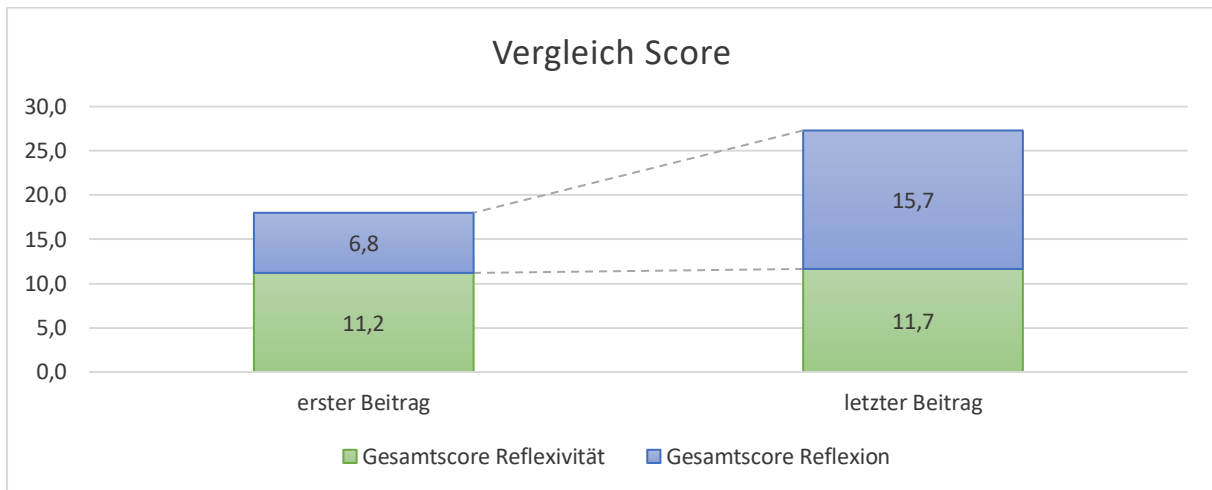
Hütemethode Verständlichkeit	Stimmt voll und ganz
Hütemethode Hilfestellung	Stimmt eher schon

P2 ist männlich und in Klasse 1. Nach den ersten drei Beiträgen erhält die Testperson Gibbs Reflexivitätszyklus, nach weiteren drei Beiträgen die Bonos Hütemethode und anschließend verfasst er noch vier Beiträge. P2 versteht zu Beginn des Projekts unter Reflektieren „über etwas nachzudenken oder etwas wiederzugeben im Sinne von *wiederspiegeln*“ (P2). In der zweiten Definition schreibt er folgendes: „Reflektieren bedeutet für mich, über eine ausgewählte Thematik genau und aus verschiedenen Perspektiven nachzudenken. Zu Beginn steht dabei immer eine Analyse der momentanen Situation, gefolgt von den möglichen Handlungsoptionen und schließlich einem Fazit in dem die zukünftigen Vorgehen ersichtlich werden“ (P2). Während er bei der ersten Definition reflektieren im Sinne von etwas Wiedergeben – vermutlich abgeleitet aus dem physikalischen Begriff – versteht, erweitert sich sein Begriffsverständnis in der zweiten Definition. Hier nimmt er die Multiperspektivität, eine Analyse, Handlungsoptionen und ein Fazit über künftiges Verhalten mit auf. Es findet sich somit sowohl Reflexion als auch Reflexivität in seinem zweiten Begriffsverständnis. Während P2 die Hütemethode als sehr verständlich und als *eher schon* hilfreich einstuft, findet er Gibbs Reflexivitätszyklus *eher schon* verständlich, aber *eher nicht* hilfreich fürs Reflektieren in den eigenen Beiträgen.

Vergleich der Beiträge zum Thema Tierschutz

Score

P2 hat bei seinem ersten Beitrag zum Thema Tierschutz einen Gesamtscore im Reflektieren von 18, der sich auf einen Reflexionsscore von 6,8 und einen Reflexivitätsscore von 11,2 aufteilt. Der zweite Tierschutz-Beitrag besitzt einen Gesamtscore von 27,4 mit einem Reflexionsscore von 15,7 und einem Reflexivitätsscore von 11,7. Insgesamt ist der Score des zweiten Beitrags um den Wert 9,4 höher als der des ersten. Auch der Reflexionsscore ist um den Wert 9,8 höher, der Reflexivitätsscore ist mit 0,5 Zuwachs fast gleich.



Analyse der Reflektierstufen

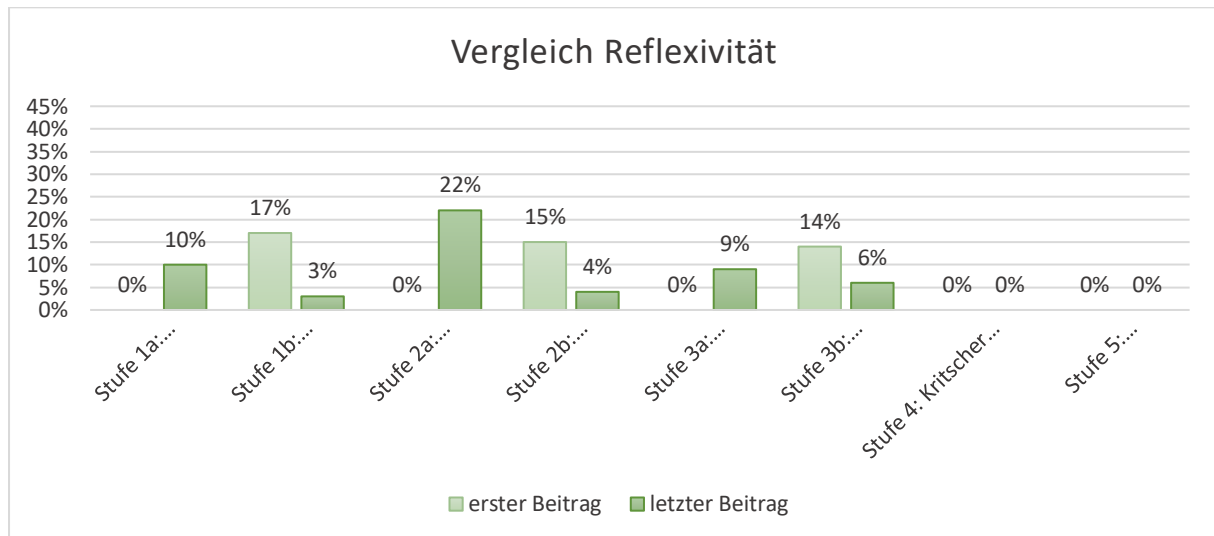
Im ersten Beitrag enthalten 46% des Textes Reflexivitätsanteile, die sich aus folgender Verteilung zusammensetzen: 17% Stufe 1b, 15% Stufe 2b und 14% Stufe 3b. P2 reflektiert im ersten Beitrag stets über seine eigene Handlung und nicht über allgemeines Handeln. Seine Sätze sind wie bei diesem Beispiel aus der Ich-Perspektive geschrieben: „Ich selbst bin kein Vegetarier, Veganer oder ähnliches und konsumiere des Öfteren Fleisch und andere tierische Produkte wie Milch, Honig oder Käse“ (P2, erster Beitrag, Reflexivitätsstufe 1b). In Stufe 3b schreibt er über sein künftiges Vorhaben und formuliert sein Ziel entsprechend: „Mein Ziel ist es also, in Zukunft mehr über mein Handeln nachzudenken und auch zu versuchen, die Menschen um mich herum für das Thema Tierschutz zu interessieren“ (P2, erster Beitrag, Reflexivitätsstufe 3b). Diese Formulierung wurde aus der Analyse des eigenen Verhaltens heraus entwickelt und formt das Fazit des ersten Beitrags. Die Stufen 1b, 2b und 3b sind im ersten Beitrag ziemlich gleich verteilt.

Der letzte Beitrag besteht zu 54% aus Reflexivitätsanteilen, die sich aus allen Stufen zwischen 1a und 3b zusammensetzen (10% 1a, 3% 1b, 22% 2a, 4% 2b, 9% 3a und 6% 3b). Insgesamt reflektiert P2 in diesem Beitrag mehr über allgemeines Verhalten als über sein eigenes Verhalten und begründet analysiert das allgemeine Verhalten hauptsächlich, wie beispielsweise in folgenden Textstellen:

Viele Menschen essen aus ethischen Gründen vegan, weil sie keine Massentierhaltung unterstützen wollen oder Tiere im Allgemeinen nicht leiden lassen wollen, es gibt jedoch noch viele andere Gründe wie das Ziel einer gesünderen Ernährung oder das Vermeiden von Antibiotika-Spritzen für Nutztiere, welche zu gefährlichen, multiresistenten Keimen führen. (P2, zehnter Beitrag, Reflexivitätsstufe 2a)

Und obwohl wir für unsere Haustiere immer nur das Beste wollen, fügen wir manchen Tieren Leid zu, sei es direkt durchs Essen ihres Fleisches und der Nutzung ihrer Produkte wie beispielsweise Wolle oder indirekt durch den Kauf und die Produktion von Waren, für die Tiere leiden oder manchmal sogar sterben mussten, obwohl das nicht immer auf den ersten Blick ersichtlich ist. (P2, zehnter Beitrag, Reflexivitätsstufe 3a)

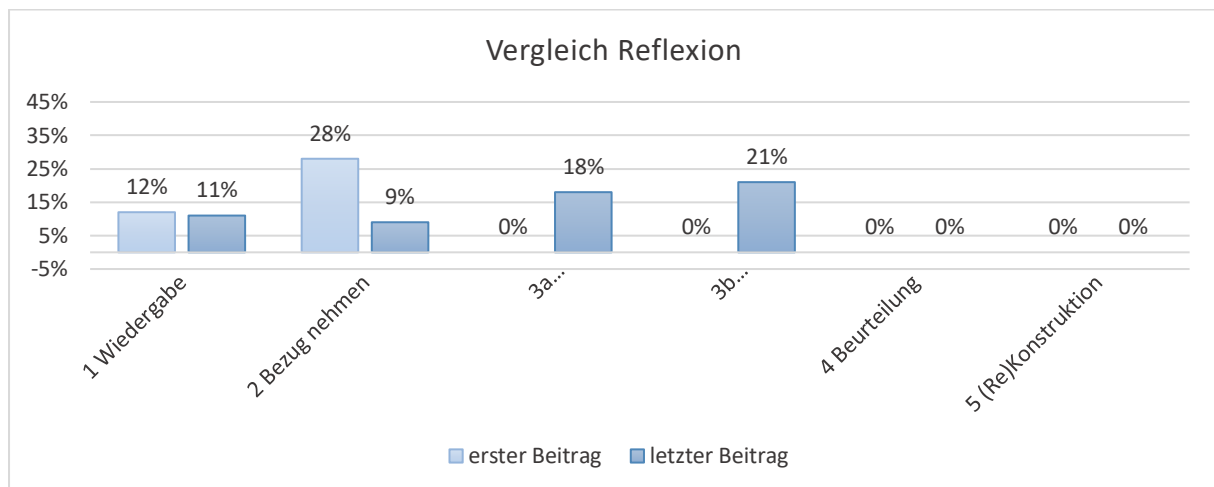
Insgesamt nimmt der Gesamtanteil der Reflexivität im Text um 8% zu, während die maximal erreichte Stufe gleichbleibt. Im letzten Beitrag wird mehr über allgemeines Verhalten reflektiert und weniger über das eigene Verhalten. Die Verteilung der Reflexivitätsstufen wird breiter, während das Niveau ähnlich bleibt.



Im ersten Beitrag von P2 bestehen 40% des Textes aus Reflexionsanteilen, die sich aus Stufe 1 mit 12% Anteil am Text und Stufe 2 mit 28% Anteil am Text zusammensetzen. Die Testperson ist folglich auf einem niedrigen Reflexionsniveau und äußert höchstens seine Meinung oder Beobachtung ohne Gründe oder tiefere Analysen, wie beispielsweise in diesem Satz: „Beim Thema Tierschutz denke ich natürlich (wie wahrscheinlich die meisten) erst einmal an Dinge wie Artenschutz, Erhalt der natürlichen Vielfalt und andere Dinge“ (P2, erster Beitrag, Reflexionsstufe 2). Der letzte Beitrag weist mit 59% Anteil am Gesamttext mehr Textstellen mit Reflexionsanteilen auf. P2 erreicht in diesem Beitrag maximal Stufe 3b mit 21% Anteil am Text, die Stufen darunter erreicht er alle (11% Stufe 1, 9% Stufe 2, 18% Stufe 3a). Die Testperson reflektiert im letzten Beitrag das Thema, indem er nicht nur seine Meinung äußert, sondern auch Gründe für sein Urteil und für Sachverhalte äußert. Folgendes Zitat verdeutlicht die Urteilsbildung mit Begründung:

Wie man vielleicht bereits aus den Zeilen herauslesen konnte, bin ich dem Einsatz von Tieren wie Mäusen für die Wissenschaft nicht negativ gegenüber eingestellt, da ich es immer noch besser finde die Experimente an Tieren durchzuführen, als Menschen an möglicherweise heilbaren Krankheiten sterben zu lassen. (P2, zehnter Beitrag, Reflexionsstufe 3b)

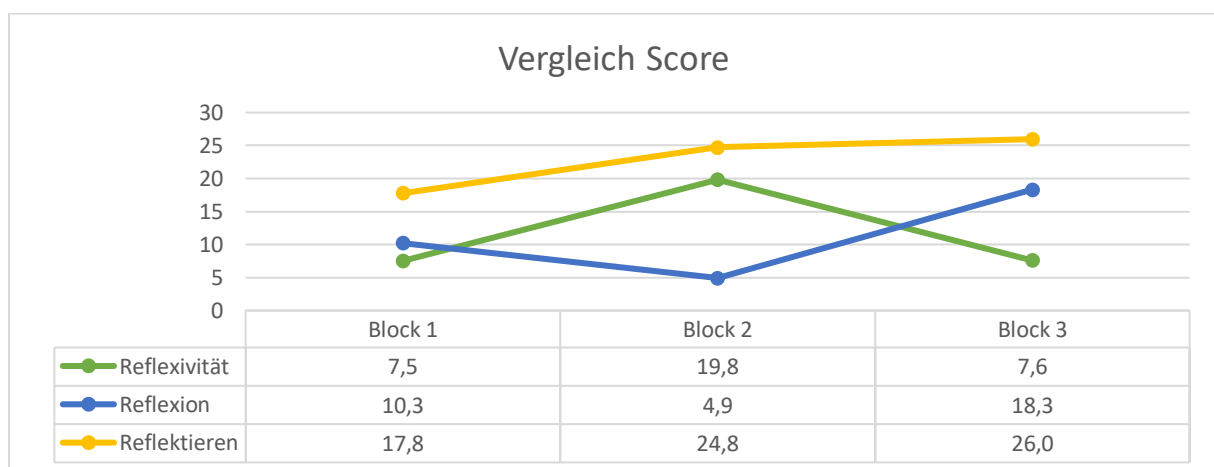
P2 verbessert sich vom ersten auf den letzten Beitrag in Hinblick auf seine Reflexionsfähigkeit. Der Gesamtanteil am Text nimmt zu, ebenfalls ist seine maximal erreichte Stufe um eine ganze Stufe höher.



Vergleich der Beiträge vor, zwischen und nach den Reflektiermethoden

Score

Der Gesamtscore im Reflektieren beträgt vor den Reflektierhilfen durchschnittlich 17,8. Der Wert setzt sich aus einem Reflexivitätsscore von 7,5 und einem Reflexionsscore von 10,3 zusammen. Nach Erhalt von Gibbs Reflexivitätszyklus steigt der Reflexivitätsscore um den Wert 12,3 auf 19,8 und der Reflexionsscore sinkt um den Wert 5,4 auf 4,9. Der Gesamtscore im Reflektieren ist im zweiten Textblock bei 24,8. Nach der Hütemethode erhöht sich der Gesamtscore nochmal leicht auf den Wert 26, der sich aus einem Reflexivitätsscore von 7,6 (nur 0,1 höher als in Block 1 vor Erhalt der Hilfen) und einem Reflexionsscore von 18,3 ergibt. Während der Reflexionsscore in Block 3 am höchsten ist, ist der Reflexivitätsscore in Block 2 am höchsten, die Werte davor und danach sind jeweils deutlich niedriger.



Analyse der Reflektierstufen

In den Beiträgen vor Erhalt der Reflektierhilfen sind durchschnittlich 35% der Textabschnitte mit Reflexivitätsanteilen. 8% des Gesamttextes ist in Stufe 1a, 7% in Stufe 1b, 15% in Stufe 2b und 6% in Stufe 3b. P2 reflektiert folglich am häufigsten über sein eigenes Verhalten, indem er Gründe dafür nennt und aus der Analyse heraus ein erweitertes Verständnis entwickelt. Allgemeines Verhalten beschreibt er lediglich. Das folgende Zitat veranschaulicht einerseits seine Überlegungen, wie er sein Verhalten ändern kann und warum es sinnvoll wäre und andererseits das sich daraus abgeleitete erweiterte Verständnis:

Indem ich versuche meinen Stromkonsum zu reduzieren und nicht mit offenem Fenster die Heizung laufen lasse und indem ich andere auch darauf aufmerksam mache, kann ich zwar noch nicht alles ändern, aber dazu beitragen etwas zu tun. In meinen Augen ist es dasselbe Prinzip wie mit der Wahlbeteiligung. Natürlich kann ich mich hinstellen und sagen, auf mich kommt es bei so vielen anderen Menschen doch nicht an, wenn aber jeder einzelne ohne Ausnahme teilnimmt, kann sich das Ergebnis doch wesentlich ändern. (P2, dritter Beitrag, Reflexivitätsstufen 2b und 3b)

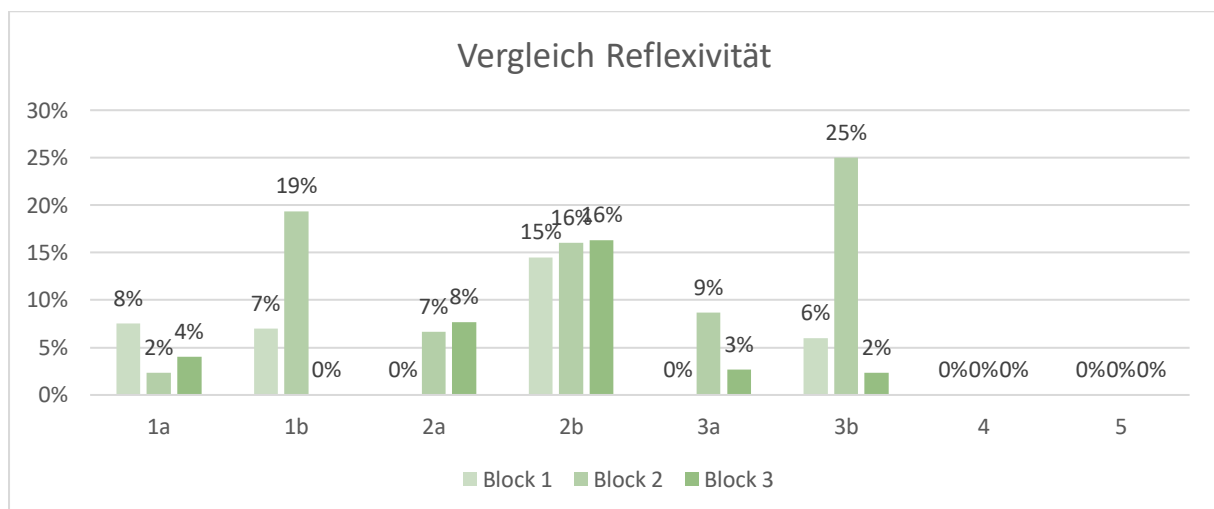
Textblock 2, der sich aus den Beiträgen nach Erhalt von Gibbs Reflexivitätszyklus zusammensetzt, weist 78% Reflexivitätsanteile im Gesamttext auf. Dabei sind alle Stufen von 1a bis 3b enthalten mit einer Verteilung von 2% in 1a, 19% in 1b, 7% in 2a, 16% in 2b, 9% in 3a und 25% in 3b. Am häufigsten reflektiert P2 in Block 2 in Stufe 3b, indem er sein eigenes Verhalten genauer analysiert und seine dabei auftretenden Gefühle berücksichtigt, wie es beispielsweise in folgendem Zitat sichtbar wird:

Alles in allem fühle ich mich nicht extrem schlecht, wenn ich über mich und mein Einkaufsverhalten nachdenke. Ich gehe nicht zu Primark, wie andere aus meiner Klasse, wir kaufen nicht nur für einen Tag ein und fahren dann sieben Mal die Woche einkaufen und wir verzichten auf unnötig mit Plastik verpackte Produkte.

In Zukunft könnte ich eventuell auf Aluminiumverpackungen und ähnliches verzichten und mir mehr faire Kleidung kaufen, solange sie nicht vollkommen überteuert ist und ich denke, dass durch ein gesteigertes Bewusstsein der gesamten Bevölkerung für Nachhaltigkeit die Nachfrage für solche Klamotten steigen und die Preise dadurch vielleicht sinken könnten, weshalb ich über diesen Blog und auch privat versuche, so viele Leute wie möglich darauf aufmerksam zu machen. (P2, sechster Beitrag, Reflexivitätsstufe 3b)

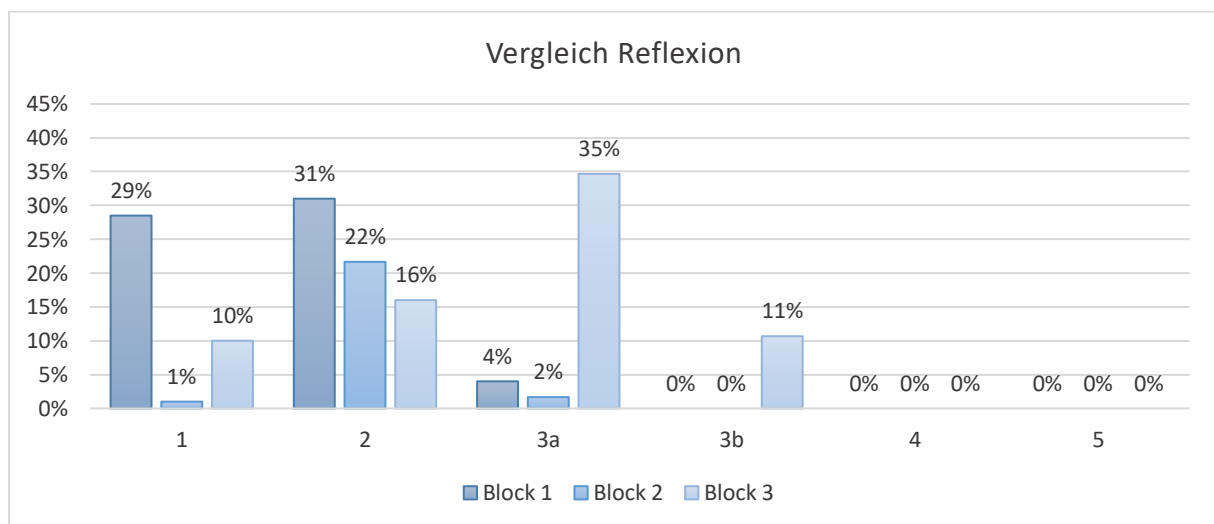
Die Testperson reflektiert wie bereits im ersten Textblock mehr über sein eigenes Verhalten und weniger über allgemeines Verhalten, allerdings steigt das Reflexivitätsniveau, auf dem er über allgemeines Handeln schreibt um zwei volle Stufen an. Er beschreibt dieses Verhalten nicht mehr nur, sondern begründet und analysiert es und schlägt Handlungsalternativen vor, wie z.B. in diesem Textausschnitt: „Von Schokolade wird zum Beispiel ein Großteil an der Elfenbeinküste unter nicht ganz so menschenwürdigen Bedingungen hergestellt, worauf man beim Einkaufen achten sollte“ (P2, fünfter Beitrag, Reflexivitätsstufe 2a).

Der dritte Textblock hat ähnlich viele Reflexivitätsanteile wie der erste Textblock mit einem Anteil am Text von 33%. Die vorkommenden Stufen sind bis auf das Nichtvorhandensein von Stufe 1b gleich der Stufen im zweiten Textblock, lediglich die Verteilung ist anders. Am häufigsten reflektiert die Testperson über seine eigene Handlung in der Reflexivitätsstufe 2b mit 16% Anteil am Text. 4% des Textes sind in Stufe 1a, 8% in Stufe 2a, 3% in Stufe 3a und 2% in Stufe 3b. Die Reflexivitätsfähigkeit in Hinblick auf die einzelnen Reflexivitätsstufen nimmt von Block 1 auf Block 2 nach Erhalt von Gibbs Reflexivitätszyklus zu und es wird intensiver über die eigene Handlung reflektiert. Außerdem geht P2 nun zusätzlich mehr auf allgemeines Verhalten ein. Im letzten Textblock, nach Erhalt der Hütemethode, nimmt die Reflexivitätsfähigkeit wieder ab, es wird vermehrt auf einem niedrigeren Niveau über die eigene Handlung reflektiert (z.B. „Ich selbst habe noch nicht sehr viele Kleidungsstücke, die komplett fair und nachhaltig hergestellt wurden, da diese häufig extrem teuer sind [...]“ (P2, achter Beitrag, Reflexivitätsstufe 2b)) und ähnlich nur mit weniger Anteil als in Textblock 2 über allgemeines Verhalten geschrieben.



Die Texte von P2 haben im ersten Block durchschnittlich 63% Reflexionsanteile, mit 29% des Textes in Stufe 1, 31% in Stufe 2 und 4% in Stufe 3a. Die Testperson reflektiert vor den Reflektierhilfen eher oberflächlich und äußert überwiegend eigene Meinungen oder Beobachtungen und begründet nur geringfügig Sachverhalte (Beispiel: „Zuerst einmal möchte ich klar festhalten, dass ich an den vom Menschen gemachten Klimawandel glaube und ich es extrem verantwortungslos von einigen Staatsoberhäupter finde, diesen zu leugnen.“ (P2, dritter Beitrag, Reflexionsstufe 2). Im zweiten Textblock sind weniger Reflexionsanteile vorhanden mit 24% Anteil am durchschnittlichen Beitragstext. Die vorkommenden Stufen sind die gleichen wie im vorherigen Textblock, am häufigsten verbleibt die Testperson in Stufe 2 mit 22%, die anderen beiden Stufe 1 und 3a treten nur kurz mit 1% und 2% Textanteil auf. Die Reflexionsfähigkeit in Hinblick auf das Vorkommen der einzelnen Reflexionsstufen bleibt gleich, P2 re-

flektiert insgesamt nur weniger über das Thema selbst im zweiten Textblock. Im dritten Textblock, nach Erhalt der Hütemethode, erreicht die Testperson maximal Stufe 3b mit 11% Anteil am Gesamttext und somit eine halbe Reflexionsstufe höher als in den vorherigen Textblöcken. Er begründet nun seine eigene Meinung und fällt Urteile, wie beispielsweise mit folgendem Wortlaut: „Diese Engstirnigkeit einiger Menschen ist sehr hinderlich, da man ohnehin bereits genug Probleme mit nachhaltiger Energie hat und Bürgerdemonstrationen da nicht sehr hilfreich sind“ (P2, neunter Beitrag, Stufe 3b). Insgesamt enthält der dritte Textblock einen durchschnittlichen Reflexionsanteil von 71% und somit den höchsten Reflexionsanteil der drei Textblöcke. Am häufigsten reflektiert P2 in Stufe 3a, indem er Begründungen und Wechselbeziehungen zu Sachverhalten äußert, wie beispielsweise in diesem Satz: „So pumpt der Konzern Nestle momentan bis zu 50.000 Liter pro Stunde aus ihren privatisierten Quellen in Äthiopien um das Wasser teuer in Europa zu verkaufen, während die ansässige Bevölkerung unter immer häufiger ausbleibenden Regenzeiten leidet“ (P2, siebter Beitrag, Reflexionsstufe 3a). Stufe 2 macht 16% des Textes aus und Stufe 1 10%. Das Reflexionsniveau hat sich folglich im dritten Textblock gesteigert und die Testperson erreicht erstmals Stufe 3b und geht somit einen Schritt weiter und bewertet nicht nur, sondern begründet seine Meinung auch, was er im ersten und zweiten Textblock nicht macht.



Testperson P3

Allgemeines zur Person

Name	P3
Klasse	1
Geschlecht	Männlich
Anzahl Beiträge	10 Beiträge
Reihenfolge Reflektiermethoden	Hüte – Gibbs
Verständnis von Reflektieren (vor Blogprojekt)	Reflektieren bedeutet für mich zb einen Text wiederzugeben. Reflektives Verhalten ist für mich, wenn sich jemand Gedanken um sein Handeln macht um ein gutes Ergebnis zu erzielen
Verständnis von Reflektieren (nach Blogprojekt)	Reflektieren ist für mich, wenn ich ehrlich über mein eigenes Verhalten nachdenke bzw. schreibe und mir dabei überlege, was ich an meinem Verhalten ändern könnte und welche Konsequenzen das hätte.
Gibbs Reflexivitätszyklus Verständlichkeit	Stimmt voll und ganz
Gibbs Reflexivitätszyklus Hilfestellung	Stimmt eher schon
Hütemethode Verständlichkeit	Stimmt voll und ganz
Hütemethode Hilfestellung	Stimmt eher schon

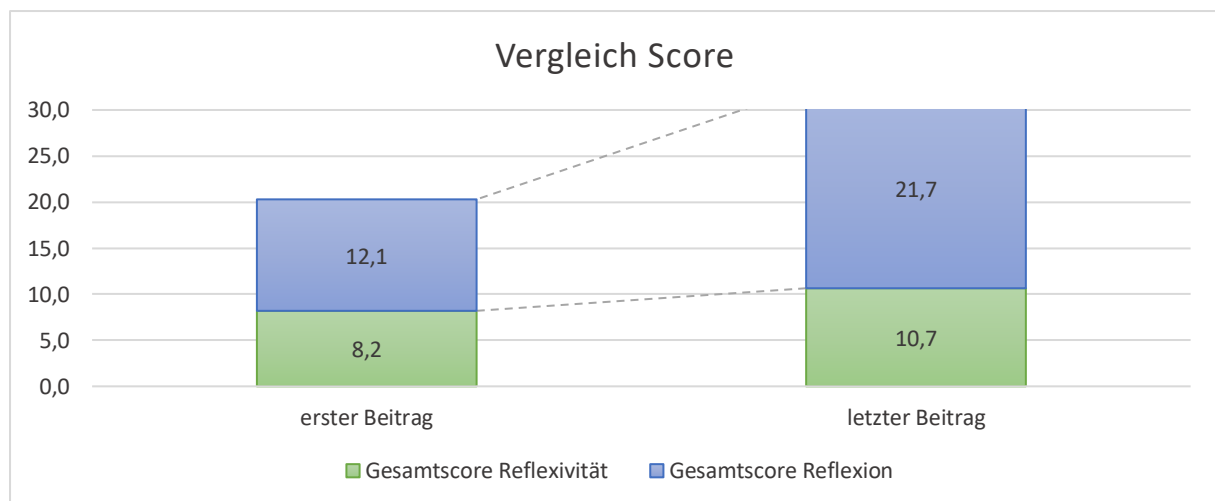
P3 ist männlich und gehört zu Klasse 1. Die Testperson hat insgesamt 10 Beiträge geschrieben und zuerst die Hütemethode und anschließend Gibbs Reflexivitätszyklus erhalten. Unter Reflektieren versteht er am Anfang des Projekts folgendes: „Reflektieren bedeutet für mich zb einen Text wiederzugeben. Reflektives Verhalten ist für mich, wenn sich jemand Gedanken um sein Handeln macht um ein gutes Ergebnis zu erzielen“ (P3). Seine zweite Definition am Ende des Projekts lautet: „Reflektieren ist für mich, wenn ich ehrlich über mein eigenes Verhalten nachdenke bzw. schreibe und mir dabei überlege, was ich an meinem Verhalten ändern könnte und welche Konsequenzen das hätte“ (P3). Zunächst nimmt er in die Definition sowohl den Aspekt der Reflexion als auch der Reflexivität mit auf, wobei er unter der Reflexion das reine Wiedergeben eines Textes versteht. In der Definition nach dem Projekt fokussiert sich die Testperson auf die Reflexivität und hierbei auf das eigene Verhalten. Er erwähnt dabei

Handlungsoptionen und das Abwägen von Konsequenzen, die sich daraus ergeben. Das Reflektieren über Inhalte nimmt er in der zweiten Definition nicht mehr mit auf. Beide Reflektierhilfen findet er *voll und ganz* verständlich und ordnet sie in *eher schon* hilfreich für das Verfassen seiner Beiträge ein.

Vergleich der Beiträge zum Thema Tierschutz

Score

Der erste Beitrag von P3 hat einen Gesamtscore im Reflektieren von 20,3 und der letzte Beitrag einen Gesamtscore im Reflektieren von 32,4. Somit hat sich dieser Score um 12,1 erhöht. Auch die Teilscores der Reflexion und Reflexivität steigen vom ersten zum letzten Beitrag an: Der Reflexivitätsscore wächst von 8,2 auf 10,7 an und der Reflexionsscore von 12,1 auf 21,7. Der Reflexionsscore hat eine deutlichere Zunahme als der Reflexivitätsscore.

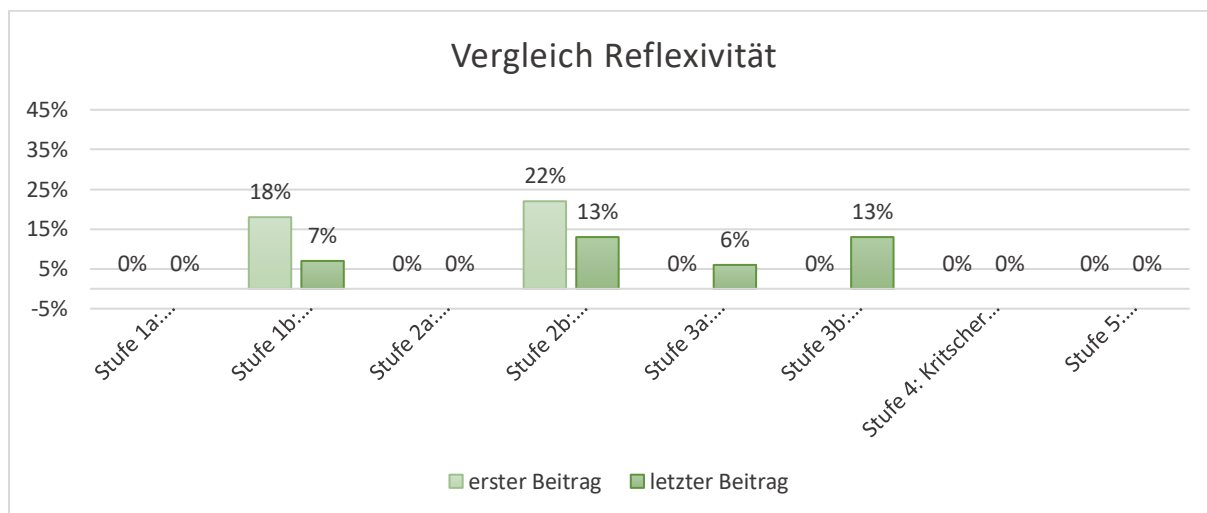


Analyse der Reflektierstufen

Im ersten Beitrag weisen 40% des Textes Reflexivitätsanteile auf, die sich aus den Stufen 1b und 2b (mit 18% und 22% Textanteil) zusammensetzen. Der letzte Beitrag, der zu 39% aus Reflexivitätsanteilen besteht, besitzt die Stufen 1b mit 7%, 2b mit 13%, 3a mit 6% und 3b mit 13% Textanteil. P3 reflektiert im ersten Beitrag nur über seine eigene Handlung und sucht nach Alternativen, z.B. „Doch eigentlich wäre es so leicht, ich müsste nur auf Prüf- und Gütesiegel achten“ (P3, erster Beitrag, Reflexivitätsstufe 2b). Im letzten Beitrag reflektiert er überwiegend auch über sein eigenes Verhalten, zusätzlich dazu auch kurz über allgemeines Verhalten. Beim eigenen Verhalten analysiert P3 im letzten Beitrag sein Verhalten genauer und begründet nicht nur einzelne Handlungsschritte, sondern berücksichtigt auch seine Emotionen diesbezüglich:

Ich bin also selbst kein Vorbild in Sachen Tierschutz. In diesen Situationen fühle ich mich meistens schlecht und auch schuldig, da man selbst irgendwie für das Leiden von anderen Lebewesen verantwortlich ist. Aber das beeinflusst mein Verhalten meist nicht über einen längeren Zeitraum. (P3, letzter Beitrag, Reflexivitätsstufe 3b)

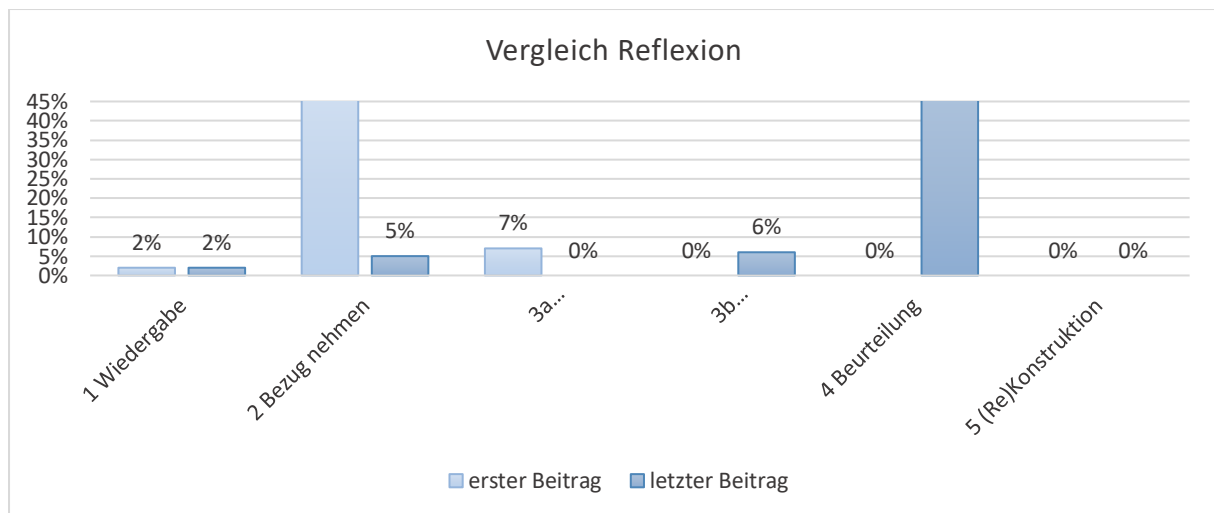
Insgesamt bleibt das Verhältnis an Reflexivitätsanteilen im ersten und letzten Beitrag gleich, die Reflexivitätsfähigkeit steigt aber dabei, wenn man sie in Abhängigkeit der auftretenden Reflexivitätsstufen betrachtet. Die höchste Stufe, die P3 im letzten Beitrag zeigt, ist um eine volle Stufe höher als im ersten Beitrag und er reflektiert nun über sein eigenes und allgemeines Verhalten.



P3 reflektiert in 58% des Textes im ersten Beitrag über den Inhalt. Mit 49% Anteil am Text ist die Stufe 2 am häufigsten vertreten. Außerdem kommen die Stufen 1 mit 2% Anteil und 3a mit 7% Anteil vor. Die Testperson äußert hauptsächlich seine Meinung und wertet Sachverhalte ohne Gründe dafür zu nennen, wie beispielsweise in diesem Satz: „Kein Tier, auch wenn es nicht vom Aussterben bedroht ist, sollte auf eine solche unwürdige Art und Weise sein Leben verlieren“ (P3, erster Beitrag, Reflexionsstufe 2). Begründungen liefert er nur für neutrale Sachverhalte ab. Im zweiten Beitrag über das Thema Tierschutz beinhalten 59% des Textes Reflexionsanteile. Darin enthalten sind die Stufen 1 (2%), 2 (5%), 3b (6%) und 4 (46%). P3 nennt nun Gründe für seine Meinungen und untersucht das Thema aus unterschiedlichen Perspektiven, erkennt Probleme und versucht Lösungen dafür zu finden. Er erreicht als höchste Stufe 4 und reflektiert am häufigsten in diesem Niveau, wie folgender Textausschnitt zeigt:

In beiden Bereichen gibt es die Sicht der Unternehmer, die durch die Tiere einen möglichst hohen Gewinn erzielen möchten und die der Tierschützer, die das ganze Thema kritischer sehen. [...] Starker Tierschutz würde sowohl Vor- als auch Nachteile für uns bedeuten. [...] Tierschutz, der unsere Freizeitgestaltung beeinflusst, wäre wahrscheinlich gegen die Meinung der meisten Menschen, wer geht schon nicht gerne mal mit seiner Familie und bestaunt Tiere, die man normalerweise nie zu Gesicht kriegen würde? [...] Eine mögliche Lösung wäre ein umfassendes und revolutionäres, weltweit gültiges Tierschutzgesetz, dass sowohl Rahmenbedingung für die Haltung von Nutztieren als auch Tiere, die zur Unterhaltung genutzt werden, bestimmt. (P3, zehnter Beitrag, Reflexionsstufe 4)

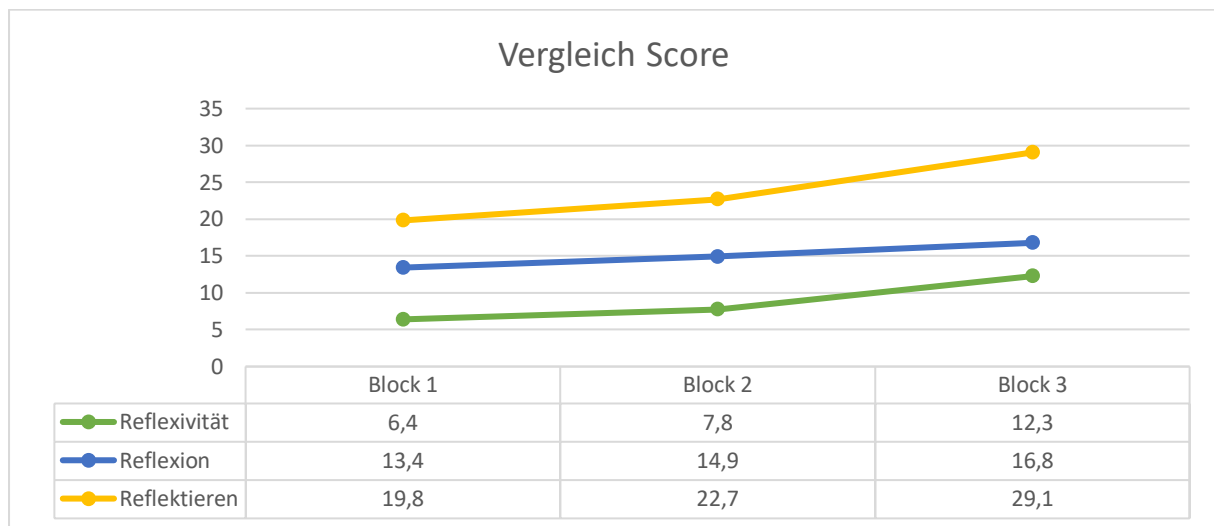
Die Reflexionsfähigkeit in Hinblick auf die einzelnen Stufen nimmt folglich vom ersten Beitrag zum letzten Beitrag zu. Die maximal erreichte Stufe von P3 steigt um eine volle Stufe von 3a auf 4. Während die Testperson im ersten Beitrag überwiegend in Stufe 2 reflektiert, reflektiert er im letzten Beitrag überwiegend in Stufe 4.



Vergleich der Beiträge vor, zwischen und nach den Reflektiermethoden

Score

Im ersten Textblock vor Erhalt der Reflektierhilfen haben die Beiträge einen durchschnittlichen Gesamtscore im Reflektieren von 19,8. Der Reflexivitätsscore beträgt dabei 6,4 und der Reflexionsscore 13,4. Nach der ersten Reflektierhilfe, in diesem Fall der Hütemethode, steigt der Gesamtscore nur leicht auf 22,7 mit einem Reflexivitätsscore von 7,8 und einem Reflexionsscore von 14,9. Der dritte Textblock nach Gibbs Reflexivitätszyklus hat einen Reflektierscore von 29,1, der aus einem Reflexivitätsscore von 12,3 und einem Reflexionsscore von 16,8 besteht. Alle drei Scores sind im dritten Textblock im Vergleich zum ersten und zweiten Textblock am höchsten. Der Reflexivitätsscore ist dabei stärker gestiegen als der Reflexionsscore.

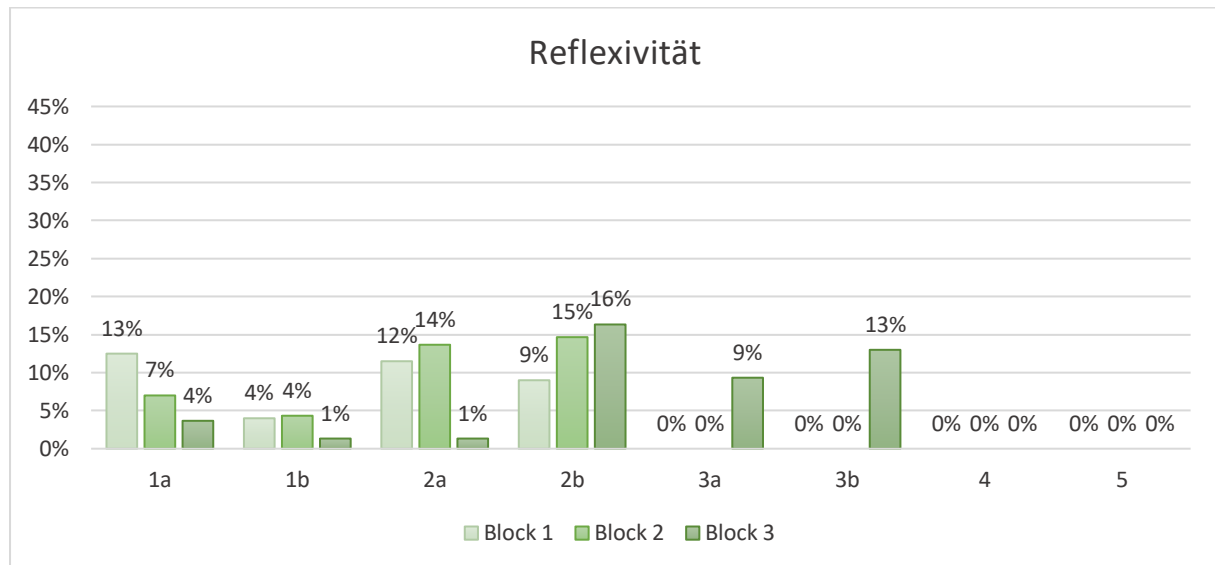


Analyse der Reflektierstufen

P3 erreicht im ersten Textblock beim Reflektieren über das Verhalten maximal Stufe 2b mit durchschnittlich 15% Anteil am Text. Außerdem reflektiert er in den Stufen 1a (13%), 1b (4%) und 2a (12%). Insgesamt weisen 37% des Textes in Block 1 Reflexivitätsanteile auf. Die Testperson schreibt über sein Verhalten und über allgemeines Verhalten, indem er Handlungen beschreibt und auch begründet. Ebenfalls schlägt er Handlungsalternativen vor. Folgendes Zitat verdeutlicht die Begründung seines eigenen Verhaltens: „Ehrlicher weißte muss ich aber auch zugeben, dass ich seit diesem Jahr an den meisten Tagen mit dem Auto, in einer Fahrgemeinschaft, zur Schule komm. Das liegt einfach an den, meiner Meinung nach unverschämte hohen, Buskosten“ (P3, zweiter Beitrag, Reflexivitätsstufe 2b). Im zweiten Textblock, bei dem die Beiträge einen durchschnittlichen Reflexivitätsanteil von 40% aufweisen, wird ebenfalls maximal Stufe 2b und alle Stufen darunter erreicht. Die Verteilung mit 7% 1a, 4% 1b, 14% 2a und 15% 2b ist ähnlich zu der des ersten Textblocks, es wird jedoch etwas mehr begründet als nur beschrieben. Der dritte Textblock hat zu 45% Reflexivitätsanteile, die sich aus den Stufen 1a bis einschließlich 3b zusammensetzen. Am häufigsten reflektiert P3 in Stufe 2b mit 16% Anteil am Text, gefolgt von Stufe 3b mit 13% Anteil und Stufe 3a mit 9% Anteil. Die Stufen 1a, 1b und 2a machen 4% und jeweils 1% des Textes aus. Die Testperson begründet häufig sein eigenes Verhalten und analysiert es in diesem Textblock auch genauer, wie folgender Textausschnitt zum Thema Strom zeigt:

Ich fühle mich dann aber doch ertappt, wenn ich mal merke, wie verschwenderisch ich damit umgehe, weswegen ich mich dann auch mal für kurze Zeit schlecht fühle. Daraus sollte ich lernen über meinen Wasserverbrauch nachzudenken, bevor ich damit etwas mache und ob es eine Möglichkeit gibt, weniger Wasser zu verbrauchen. Aus Rücksicht auf unseren Planeten will ich deswegen künftig auf jeden Fall Wasserverschwendung vermeiden! (P3, neunter Beitrag, Reflexivitätsstufe 3b)

Die Reflexivitätsfähigkeit ist vor allem vom zweiten auf den dritten Textblock gestiegen. P3 erreicht eine volle Stufe höher und reflektiert überwiegend ab dem Niveau von Stufe 2b. Vor Erhalt von Gibbs Reflexivitätszyklus findet keine Analyse von Verhalten statt, dies ändert sich danach jedoch und es wird ein Fazit über künftiges Verhalten gezogen.



Reflexionsstufen sind durchschnittlich in 71% des Textes in Block 1 enthalten. Diese teilen sich auf die Stufen 1, 2, 3 und 3b mit 17%, 48%, 3% und 4% Anteil am Text auf. In Hinblick auf die Reflexionsfähigkeit äußert P3 überwiegend seine Meinung oder wertet Sachverhalte ohne Gründe dafür zu nennen. Folgendes Zitat verdeutlicht die häufigste Reflexionsweise im ersten Block:

Meiner Meinung nach sind Elektroautos eine sehr gute und innovative Erfindung. Auch wenn die Reichweite derzeit teilweise noch recht gering und die Produktion des Stromes für die Autos noch nicht zu 100% grün ist, gehen Automobilhersteller wie Tesla, Audi, BMW und Co in die komplett richtige Richtung. Ich fände es fantastisch, wenn sich in absehbarer Zukunft ein Großteil der Bevölkerung durch komplett erneuerbare Energie fortbewegen würde. (P3, zweiter Beitrag, Reflexionsstufe 2)

Stufe 3, d.h. mit Begründungen für Sachverhalte oder Meinung erreicht er nur in Ausnahmefällen, wie beispielsweise im Satz „Ich bin trotzdem der Meinung, dass der Staat Deutschland im Stande sein sollte für alle Schüler die Buskosten zu übernehmen, da es quasi eine Investition in die Zukunft wäre“ (P3, zweiter Beitrag, Reflexionsstufe 3b).

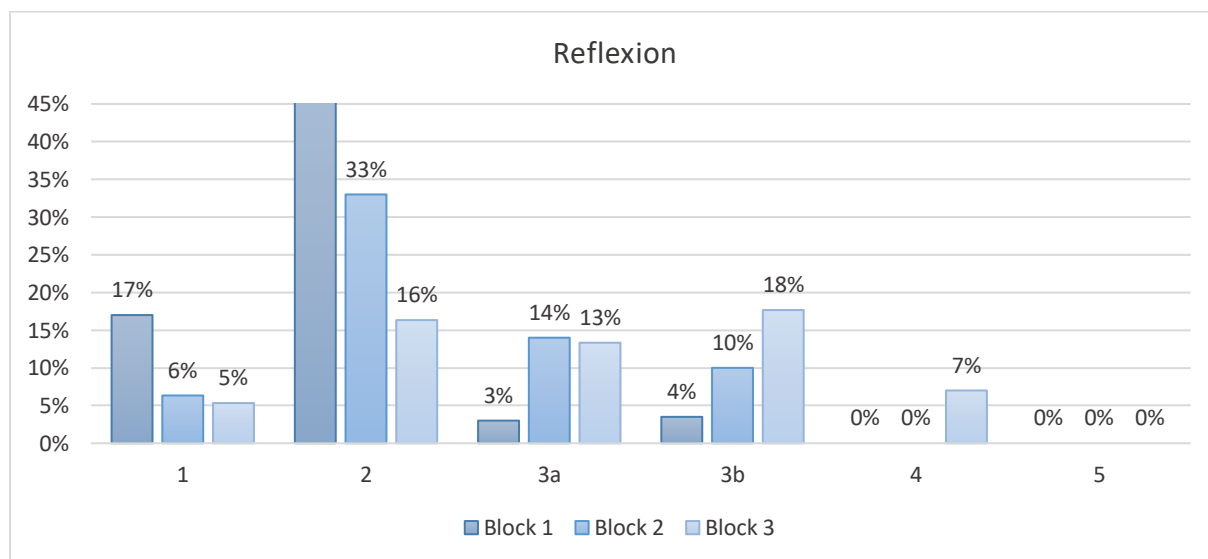
Im zweite Textblock sind durchschnittlich 63% der Beiträge mit Reflexionsanteilen. Die Stufen reichen dabei ebenfalls von Stufe 1 bis 3b, jedoch mit einer stärkeren Gewichtung auf die beiden höheren Stufen als im vorherigen Textblock. Die Stufe 1 kommt in 5%, Stufe 2 in 33%, Stufe 3a in 14% und Stufe 3b in 10% des Textes vor. Die Begründung von Sachverhalten und Meinungen stellt hier keine Ausnahme dar, sondern kommt in allen Beiträgen des zweiten Textblockes vor. Der dritte Textblock, der den höchsten Reflexionsanteil am Gesamttext mit

60% aufweist, gliedert sich in die Stufen 1 bis 4. Die höchste Stufe ist somit um eine halbe Stufe zum vorherigen Textblock gestiegen. Die Verteilung der einzelnen Stufen bleibt mit 5% Stufe 1, 16% Stufe 2, 13% Stufe 3a und 18% Stufe 3b ähnlich zum vorherigen Textblock, zusätzlich sind nun 7% des Textes der Stufe 4 zuordnungsbar:

Das muss meiner Meinung nach gestoppt werden. Und wenn nicht durch Gesetze in den betroffenen Ländern oder in Deutschland, dann durch einen Boykott unsererseits. Wenn die Marken merken, dass wir ihnen nicht alles abkaufen, werden sie etwas ändern. Aber das können wir nur gemeinsam erreichen. Und wenn nicht für die Umwelt, dann wenigstens für die Menschenwürde! (P3, achter Beitrag, Reflexionsstufe 4)

Stufe 4 wird allerdings nur in einem Beitrag erreicht, die übrigen Beiträge des dritten Textblockes haben weiterhin Stufe 3b als höchste Stufe.

Zusammenfassend kann festgehalten werden, dass sich die Reflexionsfähigkeit in Hinblick auf die Reflexionsstufen vom ersten auf den zweiten Textblock und folglich nach Erhalt der Hütemethode bezüglich der Verteilung auf die Stufen verbessert. Die Häufigkeit und maximal erreichte Stufe bleiben dabei gleich. Vom zweiten auf den dritten Textblock erhöht sich die Reflexionsfähigkeit um eine halbe Stufe von 3b auf 4, die Verteilung bleibt ähnlich der Verteilung des vorherigen Textblocks.



Testperson P4

Allgemeines zur Person

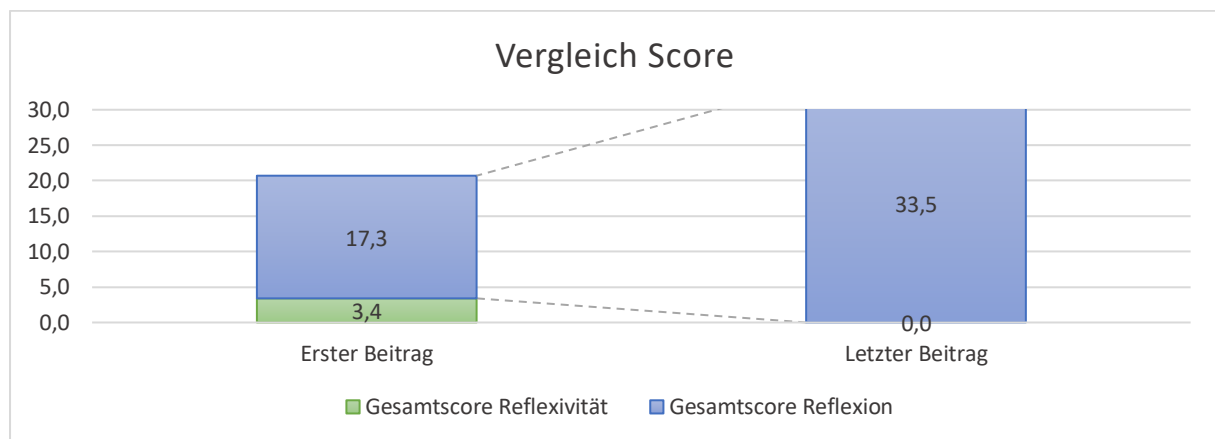
Name	P4
Klasse	1
Geschlecht	Männlich
Anzahl Beiträge	10 Beiträge
Reihenfolge Reflektiermethoden	Gibbs - Hüte
Verständnis von Reflektieren (vor Blogprojekt)	Reflektieren bedeutet für mich einerseits über gelerntes noch einmal nachzudenken, andererseits aber auch zum Beispiel ein Licht, das zurück geworfen wird.
Verständnis von Reflektieren (nach Blogprojekt)	Über mögliche Entwicklungen und Auswirkungen der Situation/des Planes nachdenken.
Gibbs Reflexivitätszyklus Verständlichkeit	Stimmt eher schon
Gibbs Reflexivitätszyklus Hilfestellung	Stimmt eher schon
Hütemethode Verständlichkeit	Stimmt eher schon
Hütemethode Hilfestellung	Stimmt eher nicht

P4 ist in Klasse 1 und männlich. Als erste Reflektierhilfe erhält er Gibbs Reflexivitätszyklus, anschließend die Hütemethode. Die Testperson hat insgesamt 10 Beiträge verfasst. Unter dem Begriff *Reflektieren* versteht er zu Beginn des Bloggens Folgendes: „Reflektieren bedeutet für mich einerseits über gelerntes noch einmal nachzudenken, andererseits aber auch zum Beispiel ein Licht, das zurück geworfen wird“ (P4). Am Ende definiert er den Begriff mit diesen Worten: „Über mögliche Entwicklungen und Auswirkungen der Situation/des Planes nachdenken“ (P4). Während er bei der ersten Definition den Aspekt der Reflexion berücksichtigt und den Begriff aus einer physikalischen Perspektive definiert, geht er bei der zweiten Definition sowohl auf die Reflexion als auch die Reflexivität ein. Beide Definitionen sind eher oberflächlich formuliert. Gibbs Reflexivitätszyklus ist für P4 *eher schon* verständlich und hilfreich, die Hütemethode ist ebenfalls *eher schon* verständlich. Für die Testpersonen ist die Hütemethode jedoch *eher keine* Hilfe für das Verfassen der Blogbeiträge.

Vergleich der Beiträge zum Thema Tierschutz

Score

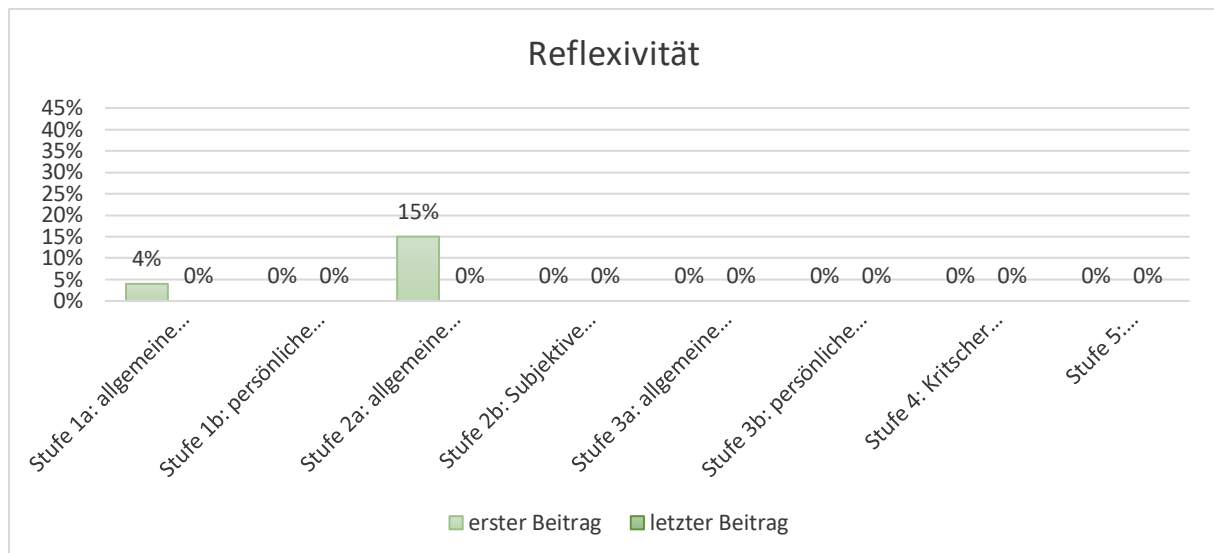
P4 erreicht beim ersten Beitrag zum Thema Tierschutz einen Gesamtscore im Reflektieren von 20,7. Der Reflexivitätsscore beträgt dabei 3,4 und der Reflexionsscore 17,3. Im letzten Beitrag zum Thema Tierschutz wird keine Reflexivitätsfähigkeit gezeigt, sodass dieser Score 0 beträgt. Der Gesamtscore im Reflektieren stimmt folglich beim letzten Beitrag mit dem Reflexionsscore überein und beträgt 33,5. Insgesamt steigt der Score im Reflektieren vom ersten auf den letzten Beitrag um den Wert 12,8. Der Reflexionsscore verbessert sich um den Wert 16,2, der Reflexivitätsscore verschlechtert sich um den Wert 3,4 auf 0.



Analyse der Reflektierstufen

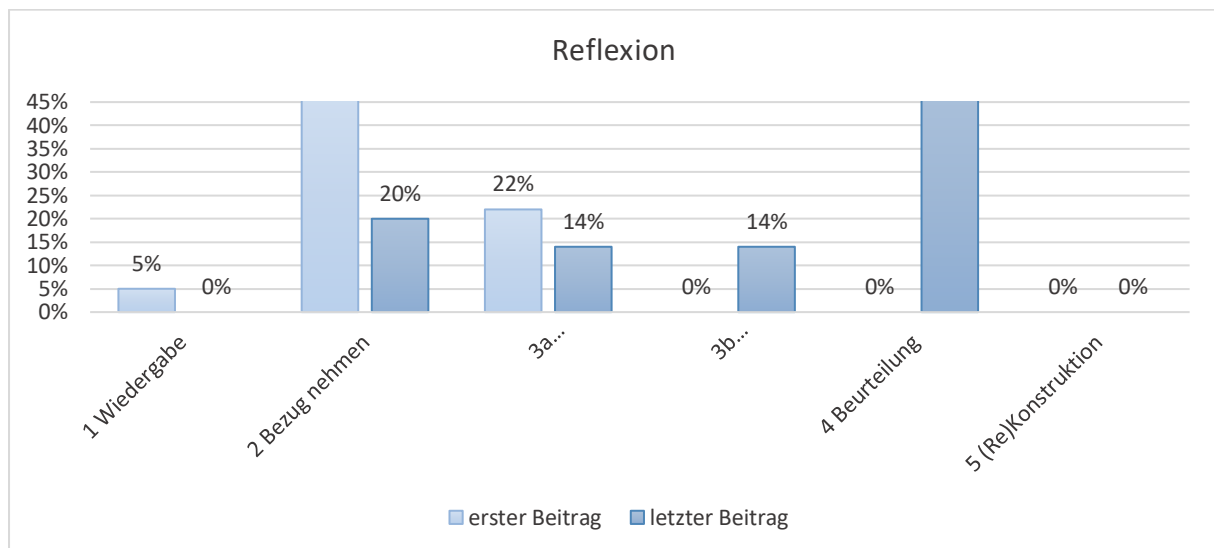
Der erste Beitrag von P4 besteht zu 19% aus Reflexivitätsanteilen. Diese sind zu 4% in der Stufe 1a und zu 15% in der Stufe 2a. Die Testperson

begründet, warum Menschen eine Doppelmoral leben mit folgenden Worten: „Einerseits wollen wir, dass es Tieren gut geht, haben selbst Haustiere, die uns viel bedeuten und finden Tierschutz eine gute Sache. Auf der anderen Seite wird im Supermarkt das Hähnchen für 2,99 € gekauft ohne groß drüber nachzudenken“ (P4, erster Beitrag, Reflexivitätsstufe 2a). Im letzten Beitrag reflektiert die Testperson nicht über sein eigenes oder allgemeines Verhalten, weshalb der Reflexivitätsanteil am Text null beträgt. Somit sinkt die maximal erreichte Stufe von 2a im ersten Beitrag um zwei volle Stufen auf null.



78% des Textes im ersten Beitrag weisen Reflexionsanteilen auf. Diese setzen sich zu 5% aus Stufe 1, zu 51% aus Stufe 2 und zu 22% aus Stufe 3a zusammen. Die Testperson begründet allgemein Sachverhalte, gibt aber keine Gründe für seine persönliche Wertung an (z.B. „Tiere werden bis zum Ende ihres kurzen Lebens unter abartigen Bedingungen gehalten“ (P4, erster Beitrag, Reflexionsstufe 2). Im letzten Beitrag gibt es insgesamt 99% Reflexionsanteile. Die gezeigte Reflexionsfähigkeit verteilt sich auf die Stufen 2 (20%), 3a (14%), 3b (14%) und 4 (51%). Die höchste Reflexionsstufe verlagert sich folglich von Stufe 3a beim ersten Beitrag auf die Stufe 4 beim letzten Beitrag. Auch die Stufe, die den meisten Anteil am Gesamttext ausmacht, verschiebt sich von Stufe 2 im ersten Beitrag auf Stufe 4 im letzten Beitrag. Die Testperson reflektiert im letzten Beitrag auf einem höheren Niveau über das Thema Tierschutz. Er beachtet unterschiedliche Perspektiven darauf (siehe Kapitel 4.2.4) und analysiert Probleme, die sich daraus ergeben. Außerdem sucht er nach Lösungen und kommt zu einem eigenen Fazit, das lautet: „Ich muss zu dem Schluss kommen, dass es in der Wirtschaft leider keine Lösung gibt, die nicht den Produktionsgrad einschränkt“ (P4, zehnter Beitrag, Reflexionsstufe 3b).

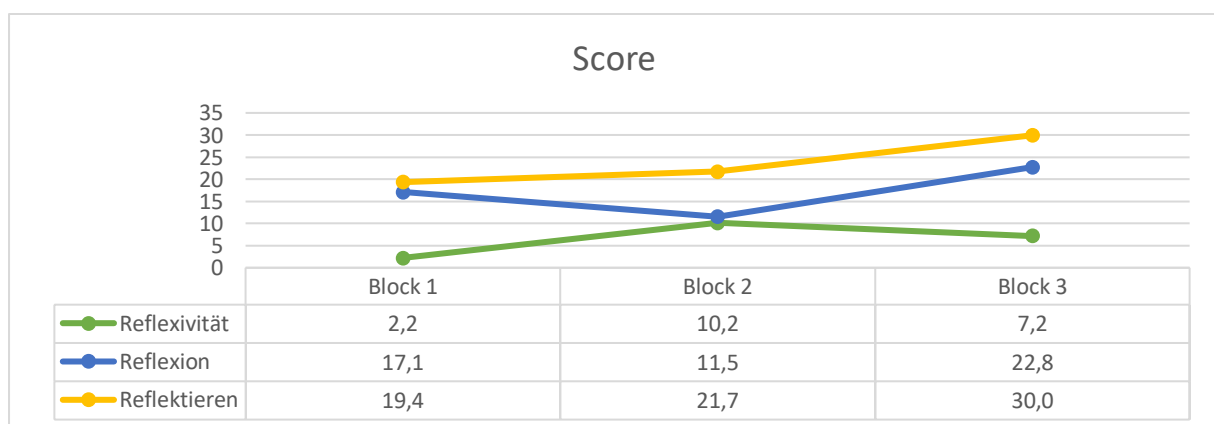
Insgesamt kann festgehalten werden, dass sich die Reflexionsfähigkeit in Hinblick auf die einzelnen Stufen vom ersten auf den letzten Beitrag um eineinhalb Stufen verbessert. Auch der Textanteil mit Reflexionsstufen steigt und die Verteilung auf die einzelnen Stufen verschiebt sich nach oben.



Vergleich der Beiträge vor, zwischen und nach den Reflektiermethoden

Score

Die Beiträge vor Erhalt der Reflektierhilfen von P4 besitzen einen durchschnittlichen Reflektierscore von 19,4, der aus einem Reflexivitätsscore von 2,2 und einem Reflexionsscore von 17,1 besteht. Textblock 2, d.h. die Beiträge nach Erhalt von Gibbs Reflexivitätszyklus haben im Schnitt einen leicht höheren Gesamtscore im Reflektieren von 21,7. Während der Reflexionsscore um den Wert 5,6 auf 11,5 sinkt, steigt der Reflexivitätsscore um den Wert 8 an. Anschließend, nach der Hütemethode, steigt der Gesamtscore auf 30, während diesmal der Reflexionsscore um fast das doppelte auf 22,8 wächst und der Reflexivitätsscore um den Wert 3 auf 7,2 fällt. In Bezug auf die Scores kann festgehalten werden, dass sich die Reflexionsfähigkeit nach der Hütemethode bessert und die Reflexivitätsfähigkeit nach Gibbs Reflexivitätszyklus. Insgesamt steigt der Gesamtscore von Block 2 auch Block 3 deutlicher als zuvor.



Analyse der Reflektierstufen

Die Beiträge des ersten Textblocks bestehen im Durchschnitt zu 12% aus Reflexivitätsanteilen. Dabei kommen die Stufen 1b mit 4%, 2a mit 5% und 2b mit 3% vor. Die Testperson reflektiert sowohl über allgemeines als auch sein persönliches Verhalten, begründet jedoch nur einzelne Handlungsschritte, ohne dass eine tiefere Analyse stattfindet (z.B.: „Mit dem Fahrrad hat es mich zweimal dorthin gezogen, da mir die Alternativen gefehlt haben und das hat lange genug gedauert.“ (P4, zweiter Beitrag, Reflexivitätsstufe 2b)). Nach Erhalt von Gibbs Reflexivitätszyklus beträgt der durchschnittliche Anteil an Reflexivität in den Beiträgen 48% und somit 36% mehr als im vorherigen Textblock. Folgende Stufen sind dabei enthalten: 1a mit 9%, 1b mit 4%, 2a mit 15%, 2b mit 11% und 3b mit 9% Anteil. Somit ist die am höchsten erreichte Stufe um eine volle Stufe höher als in Textblock 1 (von 2b auf 3b). P4 schreibt wie beim ersten Textblock sowohl über sein eigenes Verhalten als auch allgemeines Verhalten. Nun reflektiert er über seine eigene Handlung ein Niveau höher, indem er zukünftige Handlungspläne integriert:

In unserer Familie wird eigentlich schon relativ Nachhaltig eingekauft, aber was das Verfallsdatum angeht, bin ich erst mit dem Reflektieren auf den Gedanken gekommen. Ich denke, in Zukunft werde ich bewusster darauf achten, fast abgelaufene Lebensmittel einzukaufen, wenn ich weiß, dass ich es am selben Tag noch esse. (P4, sechster Beitrag, Reflexivitätsstufe 3b)

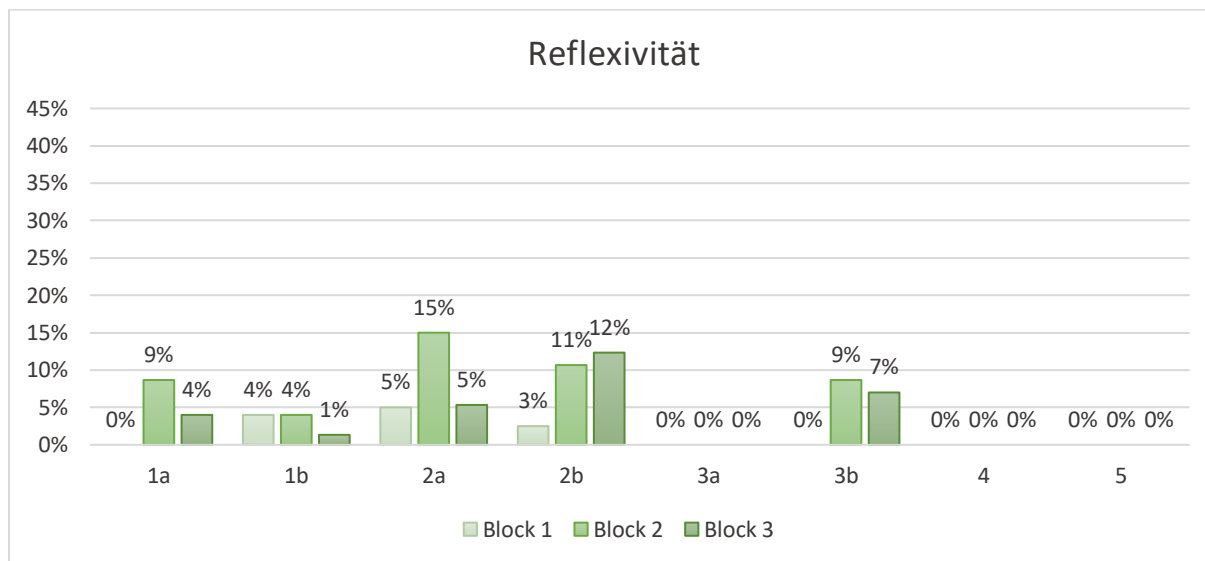
Mein Plan ist es, vor allem auf Süßigkeiten zu verzichten, aber auch möglichst nichts Verpacktes mehr zu kaufen (allein schon wegen dem Abfall), die sind nämlich die Hauptkandidaten, wenn es um unerwarteten Zuckerzusatz geht. (P4, fünfter Beitrag, Reflexivitätsstufe 3b)

Beim Reflektieren über allgemeine Handlungen bleibt er weiterhin auf dem gleichen Niveau wie in Textblock 1 und gibt lediglich Handlungsalternativen: „Wir selbst können auch dagegen steuern, indem wir darauf achten, so wenig wie möglich zuhause ablaufen zu lassen, oder beim Einkauf für's Mittagessen fast abgelaufene Lebensmittel auszuwählen“ (P4, sechster Beitrag, Reflexivitätsstufe 2a).

Der dritte Textblock weist durchschnittlich 29% Reflexivitätsanteile auf, die sich auf die gleichen Stufen wie zuvor verteilen. 1a kommt mit 4%, 1b mit 1%, 2a mit 5%, 2b mit 12% und 3b mit 7% Textanteil vor. Stufe 2a kommt seltener vor als im zweiten Textblock, die anderen Stufen sind ähnlich häufig vertreten. Auch hier schildert P4 konkret künftige Handlungsentscheidungen, die aus den Überlegungen heraus entwickelt werden, wie beispielsweise im Satz „Nach der Erkenntnis der Energiekosten für Licht werde ich in Zukunft mehr Wert auf Belichtung legen und darauf achten, nicht unnötig Lichter brennen zu lassen“ (P4, neunter Beitrag, Reflexivitätsstufe 3b).

Die Reflexivitätsfähigkeit in Hinblick auf die einzelnen Stufen verbessert sich von Textblock 1 auf Textblock 2. Die maximal erreichte Stufe liegt um eine volle Stufe höher und die Testperson reflektiert insgesamt häufiger über Handlungen zu den entsprechenden Themen, sowohl über

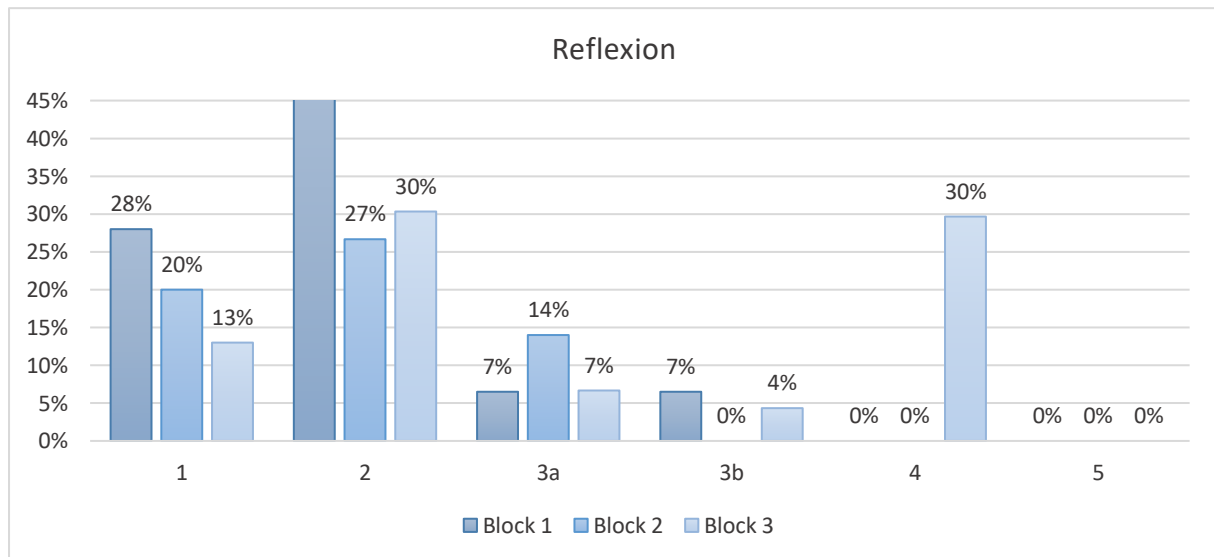
persönliche als auch allgemeine Handlungen. Während er am Anfang nur Gründe für Handlungen nennt, nimmt er ab dem zweiten Textblock die Zukunftsperspektive mit auf. Dies wird in Textblock 3 mit etwas weniger Anteil am Gesamttext fortgeführt.



Durchschnittlich 93% des Textes in Block 1 weisen Reflexionsstufen auf. Mit 51% Anteil am Text kommt Stufe 2 am häufigsten vor, gefolgt von Stufe 1 mit 28% und den Stufen 3a und 3b mit jeweils 7%. P4 äußert beim Reflektieren über das Thema oft seine eigene Meinung bzw. wertet Sachverhalte, Begründungen dafür folgen aber nicht. Auch beschreibt er viele Tatsachen ohne diese genauer zu untersuchen. Nach dem Satz „Der Klimawandel ist ein sehr langwieriger Prozess und scheint auf den ersten Blick nur geringen Einfluss zu haben.“ (P4, dritter Beitrag, Reflexionsstufe 2) folgt keine Begründung. Die Stufe 3b erreicht er, indem er das Thema mit seiner persönlichen Erfahrung vernetzt, indem er erklärt, warum sich der Klimawandel schon bemerkbar macht: „Als ich noch klein war, haben wir immer Iglus aus großen Schneehaufen gebaut. In den letzten Jahren hab ich allerdings beobachtet, dass keiner davon mehr groß genug dafür wäre. Auch ist zu sehen, dass die Vogelvelfalt nicht mehr so groß ist wie früher“ (P4, dritter Beitrag, Reflexionsstufe 3b). Der zweite Textblock weist mit 61% Anteil am Text weniger Reflexionsanteile auf. Bis auf Stufe 3b, das zuvor durch die Vernetzung des Themas mit dem eigenen Leben erreicht wurde, kommen alle Stufe aus Block 1 ebenfalls in Block 2 vor. Stufe 1 macht 20%, Stufe 2 27% und Stufe 3a 14% des Textes aus. Im dritten Textblock, nach Erhalt der Hütemethode, wird insgesamt wieder mehr über das Thema reflektiert. Durchschnittlich 83% des Textes weisen Reflexion auf, davon 13% Stufe 1, 30% Stufe 2, 7% Stufe 3a, 4% Stufe 3b und 30% Stufe 4. P4 erreicht nun beim Reflektieren eine halbe Stufe höher als zuvor und untersucht die Thematiken tiefgründiger, indem er den Problemen auf den Grund geht und nach Lösungen sucht:

Die Klospülung im Allgemeinen ist ein Knackpunkt im Wassersparen im privaten Haushalt und definitiv verbesserungswürdig. Die Familie eines alten Klassenkamerades benutzte immer gebrauchtes Spülwasser oder aufgefangenes Regenwasser zum Spülen. Außerdem gibt es auch wassersparende Klospülungen in Flugzeugen oder Zügen. Die Alternativen sind auf jeden Fall da, aber das Problem ist einfach nicht spürbar genug, um von der breiten Masse erfasst zu werden, selbst bei Nachrichten von Fernwasser, das mit Tankern in Ortschaften transportiert wird. Wirtschaftlich gesehen ist das langfristig nicht profitabel (vor allem, weil es auf Kosten des Staates passiert), aber auch moralisch ist dieser Zustand verwerflich. Man sollte meinen, dass wir mit dem Wasser klar kommen, das wir hier haben (und das ist durchaus möglich!), aber ein breitgefächelter Ausbau von Klospülungen wäre ein Schritt in die richtige Richtung. (P4, siebter Beitrag, Reflexionsstufe 4)

Dieses Reflexionsniveau zeigt P4 erstmalig nach der Hütemethode, zuvor begründet er nur seine Meinung oder Sachverhalte, ohne diese genauer zu analysieren. Während sich die Reflexion von Block 1 zu Block 2 leicht verschlechtert, da Stufe 3b nicht mehr vorkommt, nimmt die Reflexionsfähigkeit von Block 2 auf Block 3 um eine ganze Stufe zu und er reflektiert intensiver auf einem höheren Niveau.



h. Deskriptive Statistik

Deskriptive Statistik für Experimentalgruppe 1, die Gibbs Reflexivitätszyklus als erste Teilintervention erhalten						
	N	Minimum	Maximum	Mittelwert	Std,-Abwei- chung	Varianz
Gesamtscore im Reflektieren erster Beitrag	12	17,1	22,4	19,017	1,6353	2,674
Gesamtscore im Reflektieren letzten Beitrag	12	23,8	33,5	28,125	2,7180	7,388
Gesamtscore im Reflektieren Differenz letzter und erster Beitrag	12	6,5	12,8	9,108	1,7876	3,195
Reflexionsscore erster Beitrag	12	6,2	17,3	12,267	3,4778	12,095
Reflexionsscore letzten Beitrag	12	11,3	33,5	18,875	5,4681	29,900
Reflexionsscore Differenz letzter und erster Beitrag	12	2,4	16,2	6,608	3,5755	12,784
Reflexivitätsscore erster Beitrag	12	,0	11,2	6,767	3,4124	11,644
Reflexivitätsscore letzten Beitrag	12	,0	14,5	9,267	4,6091	21,244
Reflexivitätsscore Differenz letzter und erster Beitrag	12	-3,4	7,6	2,500	3,1997	10,238
Gesamtscore im Reflektieren erster Textblock	12	17,0	22,9	18,983	1,6431	2,700
Gesamtscore im Reflektieren zweiter Textblock	12	18,6	29,1	22,617	3,0614	9,372
Gesamtscore im Reflektieren dritter Textblock	12	24,1	30,0	26,442	2,2873	5,232
Gesamtscore im Reflektieren Differenz erster und zweiter Text- block	12	-,4	7,0	3,633	2,7070	7,328
Gesamtscore im Reflektieren Differenz zweiter und dritter Text- block	12	-,1	8,3	3,825	2,9093	8,464
Reflexionsscore erster Textblock	12	5,9	17,1	11,942	3,0865	9,526
Reflexionsscore zweiter Textblock	12	4,9	15,2	10,633	3,2731	10,713
Reflexionsscore dritter Textblock	12	11,0	22,8	17,050	3,2132	10,325

Reflexionsscore Differenz erster und zweiter Textblock	12	-7,5	1,6	-1,308	3,1341	9,823
Reflexionsscore Differenz zweiter und dritter Textblock	12	2,5	13,4	6,417	3,1934	10,198
Reflexivitätsscore erster Textblock	12	2,2	11,1	7,017	2,8863	8,331
Reflexivitätsscore zweiter Textblock	12	3,4	19,8	11,983	4,3541	18,958
Reflexivitätsscore dritter Textblock	12	4,9	13,9	9,392	2,5246	6,374
Reflexivitätsscore Differenz erster und zweiter Textblock	12	,0	12,3	4,967	4,0946	16,766
Reflexivitätsscore Differenz zweiter und dritter Textblock	12	-12,2	2,4	-2,592	3,9512	15,612
Maximal erreichte Reflexionsstufe erster Beitrag	12	2,0	3,5	2,958	,6201	,384
Maximal erreichte Reflexionsstufe letzter Beitrag	12	3,5	5,0	3,750	,4523	,205
Maximal erreichte Reflexionsstufe Differenz letzter und erster Beitrag	12	,0	2,0	,792	,6895	,475
Maximal erreichte Reflexivitätsstufe erster Beitrag	12	,0	3,5	2,208	,9876	,975
Maximal erreichte Reflexivitätsstufe letzter Beitrag	12	,0	3,5	2,750	1,0553	1,114
Maximal erreichte Reflexivitätsstufe Differenz letzter und erster Beitrag	12	-2,0	2,0	,542	1,0104	1,021
Maximal erreichte Reflexionsstufe erster Textblock	12	2,0	4,0	3,125	,7111	,506
Maximal erreichte Reflexionsstufe zweiter Textblock	12	2,0	4,0	3,375	,5691	,324
Maximal erreichte Reflexionsstufe dritter Textblock	12	3,5	4,0	3,708	,2575	,066
Maximal erreichte Reflexionsstufe Differenz erster und zweiter Textblock	12	-,5	1,5	,250	,5436	,295
Maximal erreichte Reflexionsstufe Differenz zweiter und dritter Textblock	12	-,5	1,5	,333	,5365	,288
Maximal erreichte Reflexivitätsstufe erster Textblock	12	1,0	3,5	2,542	,6201	,384
Maximal erreichte Reflexivitätsstufe zweiter Textblock	12	2,0	4,0	3,083	,6337	,402

Maximal erreichte Reflexivitätsstufe dritter Textblock	12	2,0	4,0	3,042	,6557	,430
Maximal erreichte Reflexivitätsstufe Differenz erster und zweiter Textblock	12	,0	1,0	,542	,4981	,248
Maximal erreichte Reflexivitätsstufe Differenz zweiter und dritter Textblock	12	-,5	,5	-,042	,3343	,112
Am häufigsten erreichte Reflexionsstufe erster Beitrag	12	1,0	2,0	1,750	,4523	,205
Am häufigsten erreichte Reflexionsstufe letzter Beitrag	12	2,0	4,0	2,458	,8382	,703
Am häufigsten erreichte Reflexionsstufe Differenz letzter und erster Beitrag	12	,0	2,0	,708	,8107	,657
Am häufigsten erreichte Reflexivitätsstufe erster Beitrag	12	,0	2,5	1,708	,7217	,521
Am häufigsten erreichte Reflexivitätsstufe letzter Beitrag	12	,0	3,5	2,333	,9847	,970
Am häufigsten erreichte Reflexivitätsstufe Differenz letzter und erster Beitrag	12	-2,0	2,0	,625	1,0687	1,142
Am häufigsten erreichte Reflexionsstufe erster Textblock	12	1,0	2,0	1,833	,3892	,152
Am häufigsten erreichte Reflexionsstufe zweiter Textblock	12	1,0	3,0	2,000	,4264	,182
Am häufigsten erreichte Reflexionsstufe dritter Textblock	12	2,0	4,0	2,250	,6216	,386
Am häufigsten erreichte Reflexionsstufe Differenz erster und zweiter Textblock	12	,0	2,0	,167	,5774	,333
Am häufigsten erreichte Reflexionsstufe Differenz zweiter und dritter Textblock	12	-1,0	2,0	,250	,7538	,568
Am häufigsten erreichte Reflexivitätsstufe erster Textblock	12	1,0	3,5	2,167	,6155	,379
Am häufigsten erreichte Reflexivitätsstufe zweiter Textblock	12	1,0	3,5	2,125	,7111	,506
Am häufigsten erreichte Reflexivitätsstufe dritter Textblock	12	1,0	2,5	1,958	,4981	,248
Am häufigsten erreichte Reflexivitätsstufe Differenz erster und zweiter Textblock	12	-1,5	1,0	-,042	,6201	,384

Am häufigsten erreichte Reflexivitätsstufe Different zweiter und dritter Textblock	12	-2,5	1,0	-,167	,9614	,924
--	----	------	-----	-------	-------	------

Deskriptive Statistik für Experimentalgruppe 2, die die Hütemethode als erste Teilintervention erhalten						
	N	Minimum	Maximum	Mittelwert	Std,-Abwei- chung	Varianz
Gesamtscore im Reflektieren erster Beitrag	12	16,9	20,5	19,175	1,0872	1,182
Gesamtscore im Reflektieren letzten Beitrag	12	25,5	34,4	30,308	3,1152	9,704
Gesamtscore im Reflektieren Differenz letzter und erster Beitrag	12	6,0	15,5	11,133	3,3358	11,128
Reflexionsscore erster Beitrag	12	5,4	15,9	11,175	3,4415	11,844
Reflexionsscore letzten Beitrag	12	9,6	28,5	19,133	5,2533	27,597
Reflexionsscore Differenz letzter und erster Beitrag	12	-4,8	16,5	7,958	6,0677	36,817
Reflexivitätsscore erster Beitrag	12	3,9	12,4	8,008	2,8843	8,319
Reflexivitätsscore letzten Beitrag	12	5,0	21,4	11,192	4,5394	20,606
Reflexivitätsscore Differenz letzter und erster Beitrag	12	-7,4	17,5	3,183	5,5925	31,276
Gesamtscore im Reflektieren erster Textblock	12	17,1	20,7	19,283	,9408	,885
Gesamtscore im Reflektieren zweiter Textblock	12	21,3	29,9	25,667	2,8030	7,857
Gesamtscore im Reflektieren dritter Textblock	12	21,3	31,0	26,092	2,7057	7,321
Gesamtscore im Reflektieren Differenz erster und zweiter Text- block	12	2,6	10,9	6,383	2,6741	7,151
Gesamtscore im Reflektieren Differenz zweiter und dritter Text- block	12	-5,1	6,4	,425	3,2558	10,600
Reflexionsscore erster Textblock	12	6,3	16,4	12,175	3,3891	11,486

Reflexionsscore zweiter Textblock	12	8,9	25,1	17,642	4,1137	16,923
Reflexionsscore dritter Textblock	12	7,7	21,3	14,750	4,0876	16,708
Reflexionsscore Differenz erster und zweiter Textblock	12	-6,2	14,3	5,467	5,3309	28,419
Reflexionsscore Differenz zweiter und dritter Textblock	12	-8,0	1,9	-2,892	3,1266	9,775
Reflexivitätsscore erster Textblock	12	2,0	12,1	7,117	3,3569	11,269
Reflexivitätsscore zweiter Textblock	12	4,8	12,4	8,050	2,2956	5,270
Reflexivitätsscore dritter Textblock	12	7,7	16,5	11,375	2,4506	6,006
Reflexivitätsscore Differenz erster und zweiter Textblock	12	-5,4	10,4	,933	4,0630	16,508
Reflexivitätsscore Differenz zweiter und dritter Textblock	12	1,2	6,5	3,325	1,4852	2,206
Maximal erreichte Reflexionsstufe erster Beitrag	12	2,0	3,5	2,792	,7217	,521
Maximal erreichte Reflexionsstufe letzter Beitrag	12	3,5	5,0	3,917	,5573	,311
Maximal erreichte Reflexionsstufe Differenz letzter und erster Beitrag	12	,0	3,0	1,125	1,0897	1,188
Maximal erreichte Reflexivitätsstufe erster Beitrag	12	2,0	2,5	2,167	,2462	,061
Maximal erreichte Reflexivitätsstufe letzter Beitrag	12	2,5	4,0	3,458	,3965	,157
Maximal erreichte Reflexivitätsstufe Differenz letzter und erster Beitrag	12	,5	2,0	1,292	,4502	,203
Maximal erreichte Reflexionsstufe erster Textblock	12	2,0	3,5	3,125	,5691	,324
Maximal erreichte Reflexionsstufe zweiter Textblock	12	3,0	5,0	3,667	,4924	,242
Maximal erreichte Reflexionsstufe dritter Textblock	12	2,0	4,0	3,500	,5222	,273
Maximal erreichte Reflexionsstufe Differenz erster und zweiter Textblock	12	-,5	1,5	,542	,6895	,475
Maximal erreichte Reflexionsstufe Differenz zweiter und dritter Textblock	12	-1,5	,5	-,167	,5774	,333

Maximal erreichte Reflexivitätsstufe erster Textblock	12	1,5	2,5	2,375	,3108	,097
Maximal erreichte Reflexivitätsstufe zweiter Textblock	12	1,0	3,5	2,750	,6908	,477
Maximal erreichte Reflexivitätsstufe dritter Textblock	12	2,0	4,0	3,375	,5691	,324
Maximal erreichte Reflexivitätsstufe Differenz erster und zweiter Textblock	12	-,5	1,0	,375	,4827	,233
Maximal erreichte Reflexivitätsstufe Differenz zweiter und dritter Textblock	12	,0	1,0	,625	,3769	,142
Am häufigsten erreichte Reflexionsstufe erster Beitrag	12	2,0	2,0	2,000	,0000	,000
Am häufigsten erreichte Reflexionsstufe letzter Beitrag	12	2,0	4,0	2,917	,9731	,947
Am häufigsten erreichte Reflexionsstufe Differenz letzter und erster Beitrag	12	,0	2,0	,917	,9731	,947
Am häufigsten erreichte Reflexivitätsstufe erster Beitrag	12	1,0	2,5	1,875	,5691	,324
Am häufigsten erreichte Reflexivitätsstufe letzter Beitrag	12	2,0	4,0	2,833	,7785	,606
Am häufigsten erreichte Reflexivitätsstufe Differenz letzter und erster Beitrag	12	-,5	2,5	,958	,9876	,975
Am häufigsten erreichte Reflexionsstufe erster Textblock	12	1,0	2,0	1,750	,4523	,205
Am häufigsten erreichte Reflexionsstufe zweiter Textblock	12	2,0	4,0	2,292	,6895	,475
Am häufigsten erreichte Reflexionsstufe dritter Textblock	12	1,0	3,5	2,250	,8394	,705
Am häufigsten erreichte Reflexionsstufe Differenz erster und zweiter Textblock	12	,0	2,0	,542	,7217	,521
Am häufigsten erreichte Reflexionsstufe Differenz zweiter und dritter Textblock	12	-3,0	1,5	-,042	1,3049	1,703
Am häufigsten erreichte Reflexivitätsstufe erster Textblock	12	1,0	2,5	1,625	,5276	,278
Am häufigsten erreichte Reflexivitätsstufe zweiter Textblock	12	1,0	3,0	1,958	,5418	,294

Am häufigsten erreichte Reflexivitätsstufe dritter Textblock	12	1,0	4,0	2,417	,8483	,720
Am häufigsten erreichte Reflexivitätsstufe Differenz erster und zweiter Textblock	12	-,5	1,5	,333	,6513	,424
Am häufigsten erreichte Reflexivitätsstufe Differenz zweiter und dritter Textblock	12	,0	2,0	,458	,6895	,475
Deskriptive Statistik für Kontrollgruppe, die keine Intervention erhält						
	N	Minimum	Maximum	Mittelwert	Std,-Abwei- chung	Varianz
Gesamtscore im Reflektieren erster Beitrag	15	14,9	22,5	19,927	2,1181	4,486
Gesamtscore im Reflektieren letzten Beitrag	15	15,4	22,2	18,640	1,9737	3,895
Gesamtscore im Reflektieren Differenz letzter und erster Beitrag	15	-3,3	1,4	-1,287	1,3627	1,857
Reflexionsscore erster Beitrag	15	6,7	22,3	13,700	4,4457	19,764
Reflexionsscore letzten Beitrag	15	6,9	22,2	12,560	4,1680	17,373
Reflexionsscore Differenz letzter und erster Beitrag	15	-11,4	1,6	-1,140	2,9914	8,948
Reflexivitätsscore erster Beitrag	15	,0	13,3	6,233	3,8249	14,630
Reflexivitätsscore letzten Beitrag	15	,0	9,3	6,087	2,9590	8,756
Reflexivitätsscore Differenz letzter und erster Beitrag	15	-4,0	8,5	-,147	2,6519	7,033
Maximal erreichte Reflexionsstufe erster Beitrag	15	2,0	3,5	2,867	,6673	,445
Maximal erreichte Reflexionsstufe letzter Beitrag	15	2,0	3,5	2,933	,6230	,388
Maximal erreichte Reflexionsstufe Differenz letzter und erster Bei- trag	15	-1,0	1,0	,067	,4577	,210

Maximal erreichte Reflexivitätsstufe erster Beitrag	15	,0	3,5	2,000	,9258	,857
Maximal erreichte Reflexivitätsstufe letzter Beitrag	15	,0	2,5	1,867	,7898	,624
Maximal erreichte Reflexivitätsstufe Differenz letzter und erster Beitrag	15	-1,5	2,0	-,133	,7432	,552
Am häufigsten erreichte Reflexionsstufe erster Beitrag	15	1,0	3,0	1,933	,4577	,210
Am häufigsten erreichte Reflexionsstufe letzter Beitrag	15	1,0	2,0	1,933	,2582	,067
Am häufigsten erreichte Reflexionsstufe Differenz letzter und erster Beitrag	15	-1,0	1,0	,000	,5345	,286
Am häufigsten erreichte Reflexivitätsstufe erster Beitrag	15	,0	3,5	1,833	,9194	,845
Am häufigsten erreichte Reflexivitätsstufe letzter Beitrag	15	,0	2,5	1,567	,7528	,567
Am häufigsten erreichte Reflexivitätsstufe Differenz letzter und erster Beitrag	15	-2,5	2,0	-,267	,9612	,924

Deskriptive Statistik für weibliche Experimentalgruppe						
	N	Minimum	Maximum	Mittelwert	Std.-Abweichung	Varianz
Gesamtscore im Reflektieren erster Beitrag	10	17,1	22,4	19,110	1,6135	2,603
Gesamtscore im Reflektieren letzten Beitrag	10	23,8	33,6	28,920	3,1176	9,720
Gesamtscore im Reflektieren Differenz letzter und erster Beitrag	10	6,5	13,7	9,810	2,4862	6,181
Reflexionsscore erster Beitrag	10	6,2	17,3	12,300	3,4506	11,907
Reflexionsscore letzten Beitrag	10	9,6	26,6	17,080	5,0336	25,337

Reflexionsscore Differenz letzter und erster Beitrag	10	-4,8	10,7	4,780	4,0466	16,375
Reflexivitätsscore erster Beitrag	10	,0	11,6	6,820	3,4409	11,840
Reflexivitätsscore letzten Beitrag	10	1,6	21,4	11,860	5,3625	28,756
Reflexivitätsscore Differenz letzter und erster Beitrag	10	1,6	17,5	5,040	4,6574	21,692
Maximal erreichte Reflexionsstufe erster Beitrag	10	2,0	3,5	3,050	,5986	,358
Maximal erreichte Reflexionsstufe letzter Beitrag	10	3,5	4,0	3,600	,2108	,044
Maximal erreichte Reflexionsstufe Differenz letzter und erster Beitrag	10	,0	1,5	,550	,5503	,303
Maximal erreichte Reflexivitätsstufe erster Beitrag	10	,0	3,5	2,150	,9144	,836
Maximal erreichte Reflexivitätsstufe letzter Beitrag	10	2,0	4,0	3,350	,5297	,281
Maximal erreichte Reflexivitätsstufe Differenz letzter und erster Beitrag	10	,0	2,0	1,200	,6325	,400
Am häufigsten erreichte Reflexionsstufe erster Beitrag	10	1,0	2,0	1,700	,4830	,233
Am häufigsten erreichte Reflexionsstufe letzter Beitrag	10	2,0	4,0	2,200	,6325	,400
Am häufigsten erreichte Reflexionsstufe Differenz letzter und erster Beitrag	10	,0	2,0	,500	,7071	,500
Am häufigsten erreichte Reflexivitätsstufe erster Beitrag	10	,0	2,5	1,550	,7619	,581
Am häufigsten erreichte Reflexivitätsstufe letzter Beitrag	10	2,0	3,5	2,800	,7149	,511
Am häufigsten erreichte Reflexivitätsstufe Differenz letzter und erster Beitrag	10	,0	2,5	1,250	,6770	,458

Deskriptive Statistik für männliche Experimentalgruppe						
	N	Mini- mum	Maxi- mum	Mittel- wert	Std,-Abwei- chung	Vari- anz
Gesamtscore im Reflektieren erster Beitrag	14	16,9	21,3	19,086	1,2133	1,472
Gesamtscore im Reflektieren letzten Beitrag	14	25,5	34,4	29,429	3,1333	9,818
Gesamtscore im Reflektieren Differenz letzter und erster Beitrag	14	6,0	15,5	10,343	3,0993	9,606
Reflexionsscore erster Beitrag	14	5,4	17,3	11,307	3,4815	12,121
Reflexionsscore letzten Beitrag	14	14,6	33,5	20,379	5,1243	26,259
Reflexionsscore Differenz letzter und erster Beitrag	14	1,2	16,5	9,071	4,8178	23,211
Reflexivitätsscore erster Beitrag	14	3,4	12,4	7,793	2,9970	8,982
Reflexivitätsscore letzten Beitrag	14	,0	12,9	9,064	3,7023	13,707
Reflexivitätsscore Differenz letzter und erster Beitrag	14	-7,4	7,6	1,271	3,7354	13,953
Maximal erreichte Reflexionsstufe erster Beitrag	14	2,0	3,5	2,750	,7003	,490
Maximal erreichte Reflexionsstufe letzter Beitrag	14	3,5	5,0	4,000	,5883	,346
Maximal erreichte Reflexionsstufe Differenz letzter und erster Beitrag	14	,0	3,0	1,250	1,0143	1,029
Maximal erreichte Reflexivitätsstufe erster Beitrag	14	1,0	3,5	2,214	,5447	,297
Maximal erreichte Reflexivitätsstufe letzter Beitrag	14	,0	4,0	2,929	1,0163	1,033
Maximal erreichte Reflexivitätsstufe Differenz letzter und erster Beitrag	14	-2,0	2,0	,714	,9550	,912

Am häufigsten erreichte Reflexionsstufe erster Beitrag	14	2,0	2,0	2,000	,0000	,000
Am häufigsten erreichte Reflexionsstufe letzter Beitrag	14	2,0	4,0	3,036	,9500	,902
Am häufigsten erreichte Reflexionsstufe Differenz letzter und erster Beitrag	14	,0	2,0	1,036	,9500	,902
Am häufigsten erreichte Reflexivitätsstufe erster Beitrag	14	1,0	2,5	1,964	,4986	,249
Am häufigsten erreichte Reflexivitätsstufe letzter Beitrag	14	,0	4,0	2,429	1,0163	1,033
Am häufigsten erreichte Reflexivitätsstufe Differenz letzter und erster Beitrag	14	-2,0	2,5	,464	1,1174	1,249

(Eidesstattliche) Versicherungen und Erklärungen

(§ 8 Satz 2 Nr. 3 PromO Fakultät)

Hiermit versichere ich eidesstattlich, dass ich die Arbeit selbstständig verfasst und keine anderen als die von mir angegebenen Quellen und Hilfsmittel benutzt habe (vgl. Art. 64 Abs. 1 Satz 6 BayHSchG).

(§ 8 Satz 2 Nr. 3 PromO Fakultät)

Hiermit erkläre ich, dass ich die Dissertation nicht bereits zur Erlangung eines akademischen Grades eingereicht habe und dass ich nicht bereits diese oder eine gleichartige Doktorprüfung endgültig nicht bestanden habe.

(§ 8 Satz 2 Nr. 4 PromO Fakultät)

Hiermit erkläre ich, dass ich Hilfe von gewerblichen Promotionsberatern bzw. –vermittlern oder ähnlichen Dienstleistern weder bisher in Anspruch genommen habe noch künftig in Anspruch nehmen werde.

(§ 8 Satz 2 Nr. 7 PromO Fakultät)

Hiermit erkläre ich mein Einverständnis, dass die elektronische Fassung der Dissertation unter Wahrung meiner Urheberrechte und des Datenschutzes einer gesonderten Überprüfung unterzogen werden kann.

(§ 8 Satz 2 Nr. 8 PromO Fakultät)

Hiermit erkläre ich mein Einverständnis, dass bei Verdacht wissenschaftlichen Fehlverhaltens Ermittlungen durch universitätsinterne Organe der wissenschaftlichen Selbstkontrolle stattfinden können.

(Ort, Datum, Unterschrift)